DOCUMENT RESUME

ED 234 814 IR 050 447

AUTHOR Larson, Ray R.
TITLE Evaluating Public Access On-Line Catalogs. Phase I: Development and Testing of Data Collection and Analysis Tools. Final Report.
INSTITUTION California Univ., Berkeley.
SPONS AGENCY Council on Library Resources, Inc., Washington, D.C.
PUB DATE Jul 81
NOTE 104p.; For related documents, see ED 231 384, ED 231 389-390, ED 231 395, and ED 231 400-404.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *College Libraries; Computer Programs; *Data Collection; Data Processing; *Evaluation Methods; Formative Evaluation; *Library Catalogs; Man Machine Systems; *Online Systems; Records (Forms); Statistical Data; Use Studies
IDENTIFIERS Library Statistics; *MELVYL; *Online Catalogs; Statistical Package for the Social Sciences; University of California

ABSTRACT

This report outlines specifications for the online transaction monitoring and questionnaire administration features of the University of California's (UC) Prototype On-Line Catalog, known as MELVYL. The development and testing of these features by the UC Division of Library Automation (DLA) is also described. This document is the final report during Phase I of a 2-year Council on Library Resources (CLR) evaluation project covering 14 online public access catalog systems in the United States and 1 in Canada. Sections of the report describe (1) MELVYL system configuration and user transaction monitoring; (2) monitor data normalization and analysis using online statistical file processing (SABS) and batch processing with the Statistical Package for the Social Sciences (SPSS); (3) online questionnaire administration; and (4) user activity during a 1981 MELVYL pretest. Preliminary comparisons of searching patterns in the online and card catalogs are included. Appendices provide an overview of the MELVYL patron interface, detailed descriptions of monitor records from the user transaction log, descriptions of the files generated by SABS and the SPSS programs, and examples of the statistical reports produced by SABS and SPSS processing. (ESR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED234814

EVALUATING PUBLIC ACCESS ON-LINE CATALOGS

PHASE I

DEVELOPMENT AND TESTING OF

DATA COLLECTION AND ANALYSIS TOOLS

FINAL REPORT TO THE

COUNCIL ON LIBRARY RESOURCES

by

Ray R. Larson

Division of Library Automation
Office of the Assistant Vice President
Library Plans and Policies
University of California
Berkeley, California
July 1981

IR050447

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Ray R. Larson
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ERIC Clearinghouse on
Information Resources
Syracuse University
School of Education
Syracuse, N. Y. 13210
(315) 423-3640

AUTHOR ABSTRACT
(To Accompany Document Submitted to ERIC/IR)

Author: Ray R. Larson

Title: EVALUATING PUBLIC ACCESS ON-LINE CATALOGS: PHASE I
DEVELOPMENT AND TESTING OF DATA COLLECTION AND ANALYSIS TOOLS.
FINAL REPORT TO THE COUNCIL ON LIBRARY RESOURCES.

Abstract (up to 200 words):

In an extensive two-year evaluation project sponsored by the Council on Library Resources, information on the features, system use, human-machine dialogue, and user satisfaction with public on-line catalogs was gathered in a nation-wide survey. This document is the final report of the University of California Division of Library Automation's (DLA) activities during Phase I of the project. It describes the specifications, development, and testing of on-line transaction monitoring and questionnaires incorporated in the UC Prototype On-Line Catalog (MELVYL). Some preliminary observations of user activity are reported from the On-Line Catalog Pre-Test, including comparisons of searching patterns in the on-line catalog with conventional card catalogs.

****
Information About Document
****

Paper presented at:

Project Report (Contract Information, if possible)

Journal Article (Complete Citation)--(Some journals do not appear in CIJE)

Other

IR 050 447

ABSTRACT

In an extensive two-year evaluation project sponsored by the Council on Library Resources, information on the features, system use, human-machine dialogue, and user satisfaction with public on-line catalogs is being gathered in fourteen library systems in the United States, and one in Canada.

This document is the final report on the Division of Library Automation's (DLA) activities during Phase I of the project. It describes the specifications, development, and testing of on-line interaction monitoring and questionnaires incorporated in the DLA Prototype On-Line Catalog (MELVYL). Some preliminary observations of user activity are reported from the On-Line Catalog Pre-Test.

EVALUATING ON-LINE CATALOGS: PROJECT OVERVIEW

The Council on Library Resources (CLR), in conjunction with the U.C. Division of Library Automation, the Research Libraries Group (RLG), OCLC, Inc., and Joseph Matthews Associates is conducting a nationwide study of on-line library catalogs. The goals of the study are:

1. To produce data and interpretations that will enable designers of public on-line catalogs to improve system interface features.

2. To produce data and interpretations that will enable libraries to improve the implementation and support services for on-line public catalogs.

3. To collect additional data that will enable libraries to extend public on-line catalog services to potential users.

The purpose of this project as a whole is to evaluate the performance of fourteen on-line catalog systems in the United States, and one in Canada. This will entail compiling a large and complex database of measurements from computer logs, and from questionnaires and interviews given to users and library staffs. Facts, inferences, and conclusions can then be drawn from the aggregated information about who uses on-line catalogs, why and how, about how effective on-line catalog services are, and what can be done to improve them.

In the process, it is expected that standards for more effective on-line catalogs will be developed, and that these will serve as guides for systems being developed or planned.

A review of the three phases of the overall project should help to place this report in context:

Phase I - Design and pretesting of specifications for project data collection instruments, including the development of a uniform questionnaire to be used by all participants.

Phase II - Data collection by each of the participants using the instruments developed in Phase I.

Phase III - Aggregate analysis by DLA for all data collected by project participants during Phase II. Phase II and Phase III will overlap in time as partial analyses are conducted as data becomes available from the participants.

PHASE I: THE DLA CONTRIBUTION

The primary goal of DLA's Phase I project has been to generate specifications for, develop, and test on-line data collection and evaluation programs that would aid in determining the needs of the University of California's user community.

The information and analyses of the data to be collected in Phase II will help guide the design, development and deployment (e.g., terminal allocation, etc.) of the production version of the on-line catalog scheduled for release in 1982.

The other main goal in developing the user interaction monitoring and evaluation programs described in this report is to advance the state of knowledge concerning the behavior and requirements of on-line catalog users. This information should be of interest to library staff and administrators who are developing, or considering, public on-line catalogs, as well as to the library research community.

CONTENTS

ABSTRACT ii

EVALUATING ON-LINE CATALOGS: PROJECT OVERVIEW iii

PHASE I: THE DLA CONTRIBUTION iv

Section page

1. INTRODUCTION 1

2. MELVYL: THE UNIVERSITY OF CALIFORNIA ON-LINE UNION CATALOG 3

MELVYL System Configuration 4
Hardware 4
Software 4
User Transaction Monitoring 7
Patron Interface Logs 7
ADABAS Transaction Logs 10
Batch Evaluation Programs 11

3. MONITOR DATA NORMALIZATION AND ANALYSIS 13

SABS - ON-LINE CATALOG STATISTICAL FILE PROCESSING 13
SCAN Mode 14
NORM Mode 15
SABS Processing Sequence 16
SPSS PROGRAMS 16
PROJECTED EXTENSIONS TO SABS 17
ADABAS/NFAC Linking of Log Records 17
Session Translation Scan 17
On-Line Questionaires 18
Comments Command 18
Concordances of Command Text 19
DISTRIBUTION OF MELVYL STATISTICAL REPORTS 19
Daily Runs 19
Weekly Runs 20
Monthly Runs 20
Internal Report Distribution 20
External Report Distribution 20

4. ON-LINE QUESTIONNAIRE ADMINISTRATION 21

5. MELVYL PRETEST: SOME EARLY RESULTS 24

THE AVERAGE USER SESSION 24
DISPLAYS 25
SEARCHING PATTERNS IN THE ON-LINE CATALOG 26
OBSERVATIONS AND CONCLUSIONS 27

BIBLIOGRAPHY 29

Appendix page

A. THE MELVYL PATRON INTERFACE 30

The Patron Interface 30
Look-Up Mode 30
Tutorial Mode 30
Command Mode 31
Access points 31
Primary keyword search indexes 31
Qualifyirg (secondary) search indexes 32
Fixed format control numbers 32
Searching in the Prototype On-Line Catalog 33
Boolean searching 33
Keyword modifications 34
Display formats 34
Review display 34
Brief display 35
Long display 35
MARC display 35
Tags display 35
Browsing 35
HELP 36

B. USER TRANSACTION LOG: CODES AND CONTENTS 37

VARIABLES COMMON TO ALL RECORDS 37
QCONTRL SUBSYSTEM RECORDS 39
QDISPLA SUBSYSTEM RECORDS 41
QHELP SUBSYSTEM RECORDS 42
QLOOKUP SUBSYSTEM RECORDS 42
QSEARCH SUBSYSTEM RECORDS 43

C. SABS OUTPUT FILES 44

NCDQ - ACTIVITY SUMMARY FILE 44
NCDR - DISPLAY SUMMARY FILE 45
NCDS - SESSION SUMMARY FILE 48
NCDT - SEARCH SUMMARY FILE 51
NCDU - COMMAND TRANSITION FILE 54
NCDV - SEARCH TEXT FILE 55
NCDW - ERROR TEXT FILE 56
NCDX - SCREEN TRANSITION FILE 57

NCDY - SUBSYSTEM TRANSITION FILE 58
NCDZ - SELECTIVE RECORD FILE 59
NFAO - HELP SUMMARY FILE 59
NFAP - ERROR SUMMARY FILE 60
NFAQ - BROWSE SUMMARY FILE 61
NFAR - BROWSE TEXT FILE 62

D. SPSS PROGRAMS 64

ACTSUM : ACTIVITY SUMMARY PROCESSING 64
BRWSREC : BROWSE SUMMARY PROCESSING 64
COMMTRNS : COMMAND TRANSITION PROCESSING 65
DISPSUM : DISPLAY SUMMARY PROCESSING 65
ERRREC : ERROR SUMMARY PROCESSING 65
HELPREC : HELP SUMMARY PROCESSING 66
NEWADA : ADABAS LOG FILE PROCESSING 66
SCRNTRNS : SCREEN TRANSITION PROCESSING 66
SESSUM : SESSION SUMMARY PROCESSING 66
SRCHSUM : SEARCH SUMMARY PROCESSING 67

E. REPORT EXAMPLES 68

LIST OF FIGURES

Figure page

1. MELVYL System Configuration 6
2. Transaction Monitor Overview 9
3. Questionnaire Subsystem Logic Flow 23

LIST OF TABLES

Table page

1. Files currently generated by SABS 15

Section 1

INTRODUCTION

In order for an information system, such as an on-line catalog, to provide effective and efficient service to its intended users, two things are required: first, the system must be flexible enough to change over time in accordance with the needs of its users, and second, some way must be found to determine those needs, providing feedback to the system design and development process.

In designing public access on-line catalogs, this process of feedback and refinement is even more critical than in other types of information systems, both because of the relative 'youth' of on-line catalogs (with all of the unknowns that implies) and the wide range in experience of potential users.

Most on-line information retrieval systems (such as DIALOG, ORBIT, OCLC, and RLIN) have been developed with the implicit assumption that the users will be a well defined group of persons who may be trained in the system's use.

That assumption is not tenable with public access on-line catalogs, so special efforts must be made to develop a system that is 'user-friendly' and to determine who is using the catalog, how they are using it, and what their needs and expectations are.

Throughout the period of planning and design for the University of California On-Line Union Catalog, the need for evaluation and feedback to the design process was recognized and acted upon. Early in the design phase it was decided that the evaluation process would need to incorporate both observations of actual user activity at the terminal and questionnaires to determine the motivations and needs of users. In order to make such data collection cost-effective and efficient DLA incorporated a software transaction monitor into the 'Patron Interface' and developed a sub-system to administer questionnaires on-line.

Transaction monitors have been used in a number of studies[1] of on-line information retrieval systems to record and observe user activity. Monitoring is particularly well suited to studies of public access on-line catalogs, where the user's privacy is protected by anonymity (there are no identifying 'account numbers'.)

The philosophy adopted in developing the monitoring system for the U.C. On-Line Union Catalog has been to capture enough information to be able to fully reconstruct any user session, including all commands, errors, and system responses, without intruding on the user, or degrading system performance.

This report describes: (1) the user transaction monitor developed for the catalog, (2) the on-line questionnaire administration sub-system of the catalog currently under development, and (3) presents some early observations of catalog use from the On-Line Catalog Pre-Test.

The body of the report provides a general discussion of the above system, and the appendices provide detailed descriptions of monitor records, summary records from the batch analysis programs, and examples of the statistical reports available from the system.

---

[1] See Penniman and Dominick for an excellent review and examination of monitoring.

Section 2

MELVYL: THE UNIVERSITY OF CALIFORNIA ON-LINE UNION CATALOG

The University of California on-line catalog, now in its pre-test and due for its prototype phase in a month, is a product of years of planning and development by the University of California. Over the last several years DLA staff have designed and developed programs to merge the catalog records of library holdings on all nine UC campuses, producing a very successful union list of serials and a union list of monographs, both on microfiche.

The database developed for the microfiche catalog of monographs (which contains all records of 1976-and-later imprints as well as many additional records from retrospective conversion projects on each campus) is now being used for the new on-line catalog. In one year this database expanded fivefold from 148,000 records of books to well over 700,000 records, and in 1982, when the production version of the on-line catalog is to be released, the database is expected to contain approximately two million records.

Throughout this development, University Librarians and their staffs, as well as the personnel in DLA's parent organization, the Office of the Assistant Vice President--Library Plans and Policies, have played a continuing role in planning, testing, and evaluating at every step of the way.

With public on-line catalog development still in its infancy, the University of California has already made a major contribution to the field in its years of analysis and careful development of a catalog with which, it is hoped, library users of any level of experience can feel comfortable.

2.1 MELVYL SYSTEM CONFIGURATION

The MELVYL system has been designed with the needs of a diverse user population in mind. Two modes of interaction are now in use: one for the beginning or occasional user and one for the experienced user. A third mode designed to help beginners become experienced users is under development. Care has been taken to avoid computer (and library) jargon wherever possible. The extensive 'Help' facility of the catalog is always available to users, providing not only explanations of all commands, but also the unique ability to diagnose what has occurred in a user's session and to provide a list of options for proceeding, or getting more help.

This 'Patron Interface' and the search capabilities of MELVYL are more fully described in Appendix A of this report.

The remainder of this section provides an overview of the MELVYL system configuration, and describes the monitoring capabilities incorporated into the system.

2.1.1 Hardware

The DLA computing facility consists of twin Magnuson M80/4 CPUs running under OS/360 and ASP with 8 Mbytes of real core each. On-line storage for the facility consists of 20 IBM 3350 compatible CDC disk drives (625 Mbytes each). Telecommunications hardware includes a CCI telecommunications controller supporting a statewide communications network with high speed data lines to remote terminals on all nine U.C. campuses.

2.1.2 Software

The MELVYL on-line catalog runs as three 'jobs' with asynchronous communication on a dedicated CPU. These jobs are telecommunications, the user or patron interface, and data base access and retrieval.

1. Telecommunications: The MELVYL system uses the MILTEN telecommunications interface developed at Stanford University. MILTEN mediates all transfer of data to and from the remote terminals and the patron interface, via an assembly language interface developed at DLA which permits MILTEN to communicate directly with running PL/1 programs.

2. Patron Interface: The patron interface software is written in PL/1 and was developed at DLA. The patron interface software keeps track of every terminal connected to the system, parses incoming messages (queries, etc.) from terminals, evokes appropriate subroutines (for retrieval from the data base, display formatting, error diagnosis and help messages, etc.) and provides an asynchonous interface to the database management component.

3. Data Base Access and Retrieval: The MELVYL system uses the ADABAS (Adaptable DAta BAse System) D.B.M.S. package developed and marketed by Software A.G. to manage the catalog data base.

Figure 1 provides a graphic depiction of the MELVYL system configuration.

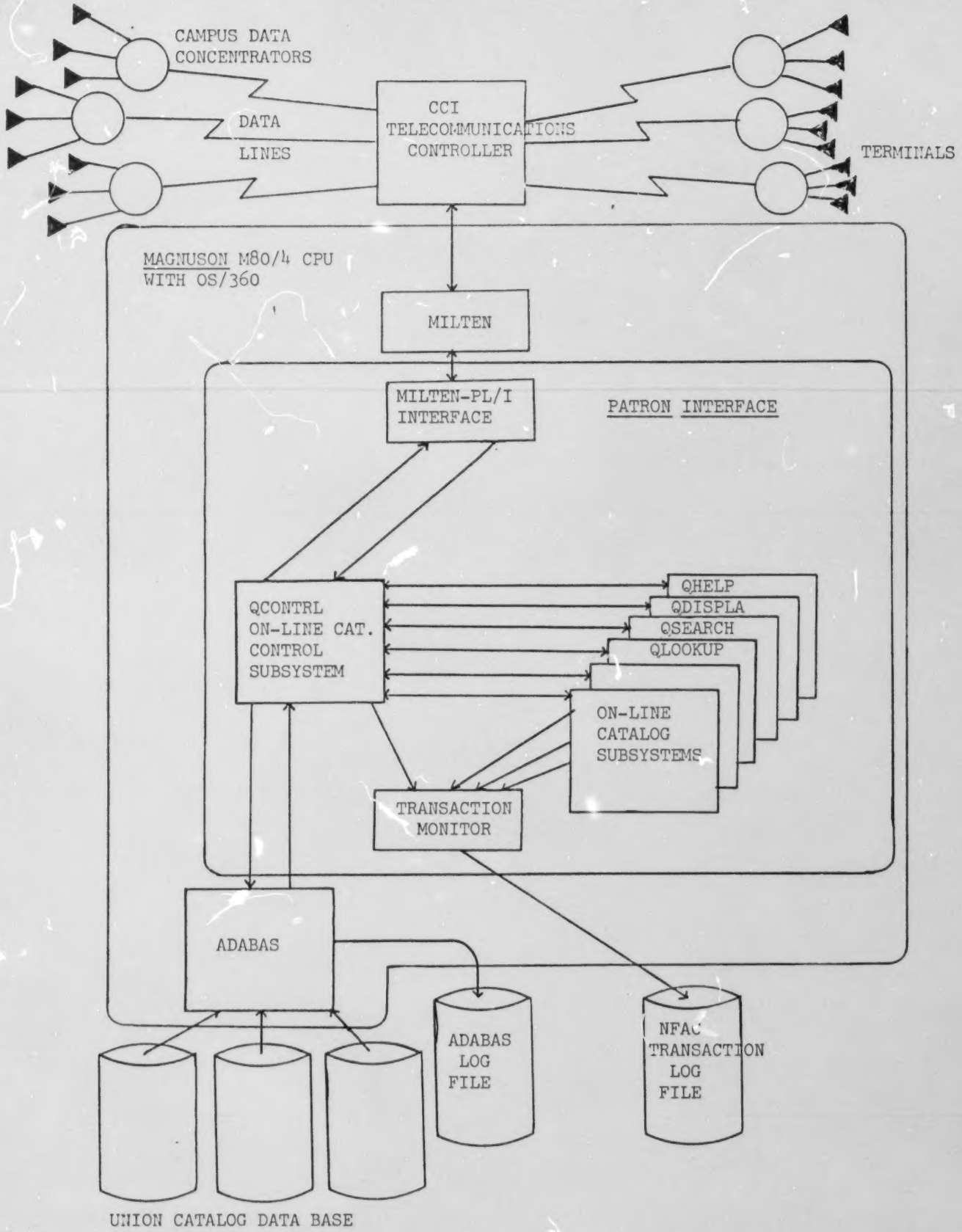


Figure 1. MELVYL System Configuration

2.2 USER TRANSACTION MONITORING

Both the patron interface and data base components of the MELVYL system generate logs of system activity.

2.2.1 Patron Interface Logs

During normal operation of the MELVYL system a software monitor incorporated into the patron interface program generates log records of virtually every significant event taking place.

Log records are constructed with variable formats, depending on the event generating the record. A certain portion of each record, however, is fixed. This fixed portion includes the following:

1. A 'Date-Time' stamp giving the date and time of day the record was logged.

2. The session number associated with a particular event (session numbers are assigned sequentially by the system at logon and are associated with all records of transactions with a particular terminal until log-off, time-out, or disconnection.)

3. The terminal address associated with the session. (Terminal addresses may be decoded to determine actual location of a particular terminal.)

4. A further time stamp showing the time of day the last command was input for a particular session.

5. A code identifying the contents of the variable portion of the log record. (These codes and their associated record contents are described in Appendix B.)

Log records are usually generated as the last step of processing immediately before a response is transmitted to the user. Records are written asynchronously to the statistical log file called NFAC.[2] (Asynchronous output means that records are not written out to the log file immediately at the moment they are completed, rather a large buffer in memory accumulates the records, which are written out when

---

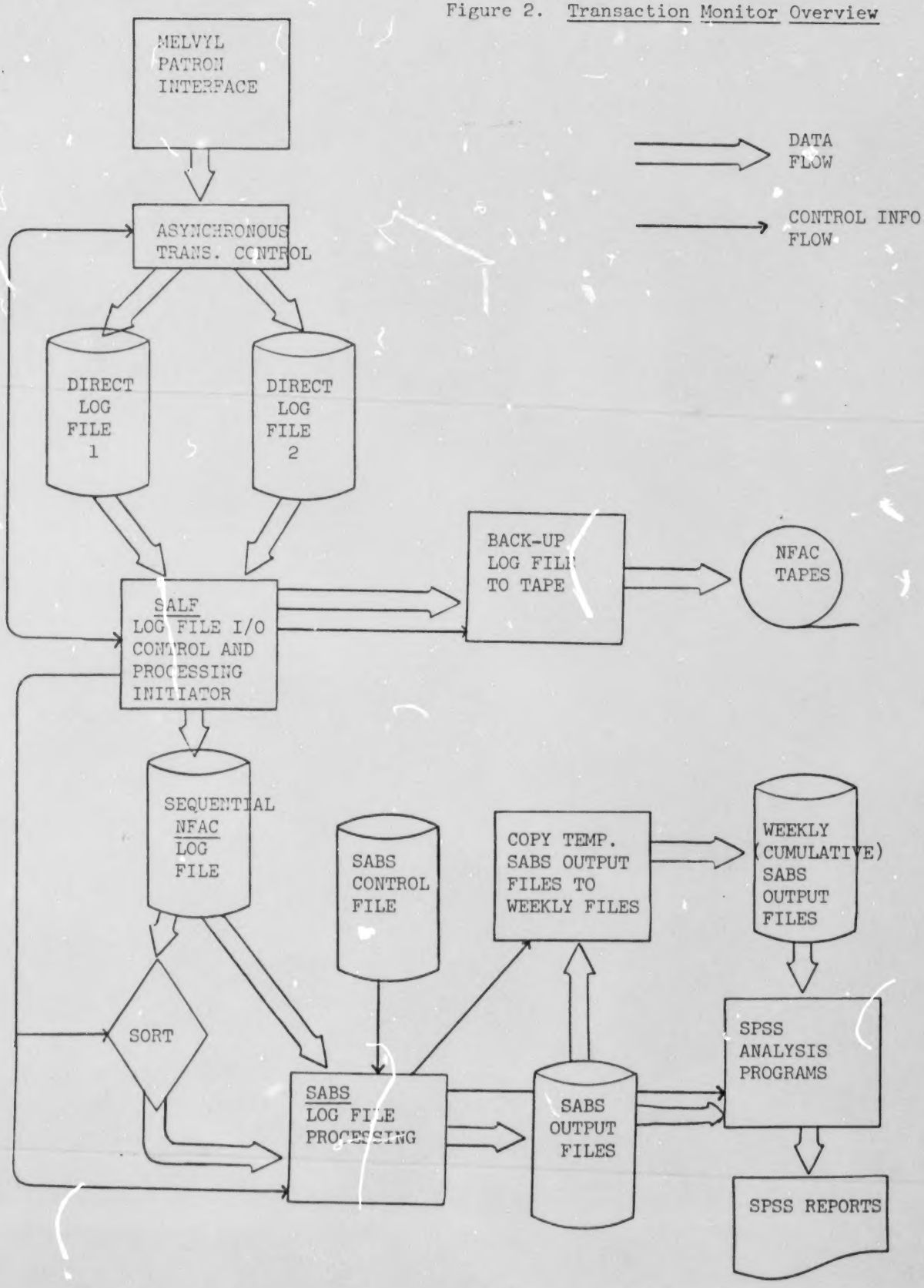
[2] The four letter names used for files and programs have no intrinsic meaning. They are just conventions used in the DLA structured programming environment to avoid having different programs, files, or data structures with the same names.

the catalog has 'free time.') The asynchronous output prevents the monitoring facilities from significantly degrading response time for the users of the system.

Figure 2 depicts the transaction monitor asynchronous output processing.

Records generated by the various subsystems of the online catalog's patron interface are passed to a subroutine (called QFN0005) that fills in the fixed record identification and time-stamp information and stores the information in a buffer area in main memory. Another subroutine keeps tabs on the status of the buffer, and when it is almost full checks for periods of low user activity and writes the contents of the buffer to a direct access file. There are two of these direct access files, but only one is used to accumulate log records at any one time. When the current file is filled (representing about one day's accumulation of log records) the subroutine switches all writing to the second file and initiates the processing of the first (full) file. (The processing sequence is described in Section 3.) Once processing is completed the first file is freed to be rewritten when the second file becomes full.

Figure 2. Transaction Monitor Overview



2.2.2 ADABAS Transaction Logs

The ADABAS D.B.M.S. generates log records of every 'call' to ADABAS from the patron interface program. Output from the ADABAS log is in a fixed format and includes the following data:

1. Time: The time of day the call was received.

2. User ID: This is the same value as the session number in the patron interface monitor records.

3. CID: CID (Command ID) is a number assigned by ADABAS to link searches requiring multiple calls.

4. Command Code: A code summarizing the type of action to be taken by ADABAS. (E.g., read inverted file, get a record from storage, perform Boolean operations between lists of records, etc.)

5. File Number: The file acted on by the call (nine different files are used in the MELVYL system).

6. ISN: (Internal Sequence Number) The record number of a record being read - when needed.

7. Response Code: A code indicating normal completion of the ADABAS call, or errors.

8. ECB Count: The number of stacked ADABAS calls pending.

9. Thread: The multiprocessing 'thread' number used for this call.

10. CPU Time: The number of CPU milliseconds used to resolve the call.

11. Associator EXCPs: Number of I/O operations to the ADABAS inverted files.

12. Data Storage EXCPs: Number of I/O operations to ADABAS data files.

13. Work Data Set EXCPs: Number of I/O operations to temporary working storage.

ADABAS uses an asynchronous output system similar to that described for the user transaction. Various utility programs (provided with the ADABAS software package) empty the dual log files and convert the internal log record storage format to fixed length records in a print format file. These 'report' format records are run through a PL/1 program to re-

move page headings and other extraneous material from the print file. The records may then be processed by SPSS.

The log files generated by the transaction monitor and ADABAS may be linked on time and user ID variables. Generally each user command will result in two or more patron interface monitor records and as many as a dozen ADABAS transaction records (depending on the complexity and type of the user command).

2.2.3 Batch Evaluation Programs

A batch program (initiated by the asynchronous transaction log output subroutine) is used to condense and normalize the daily log record files, producing a variety of output files containing summary and 'event transition' records in a format for input to SPSS (Statistical Package for the Social Sciences).

The following are some of the output reports and files generated by the batch program (SABS).

1. System Activity Report: This report shows the overall activity during the operation of the catalog system in histogram format for specified time intervals. For each interval data such as the number of active terminals, number of logons, number of logoff, number of commands received, number of screens transmitted, average response time, and program idle time are displayed graphically.

2. Session Summary: One summary record is generated for each user session containing the session length, terminal location, number and types of commands issued during the session, etc.

3. Search Summary: One summary record is generated for each search, including the number and types of indexes accessed, Boolean operators used, keyword modifications used (i.e., truncation, phrases, etc.), number of items retrieved, response time, etc. (Through linking with the ADABAS transaction logs, the CPU time and number and types of EXCPs required to resolve the search will be included.)

4. Display Summary: One summary record is generated for each display screen sent to the user. Each summary record includes the number of records displayed, display format, response time from the display command to the actual display, user 'think time' from the display to the next command, etc.

5. Command Transition Records: A record is generated for each new command issued. Codes are included giving the type of command for both the current and previous commands issued. The elapsed time between the two commands is also recorded.

This is not a complete list of output records from SABS. The program currently in operation is by no means a 'final version' but will continue to change and expand, accomodating new developments in the on-line catalog.

The operation of the SABS program is discussed more fully in the following section. The format and contents of the output records are described in Appendix C.

Section 3

MONITOR DATA NORMALIZATION AND ANALYSIS

During normal running of the catalog, a statistical log file, NFAC, is produced. It provides a variety of data concerning the activity of the Catalog and its users.

The production of the NFAC file is discussed in section 2, and details of record formats and codes are given in Appendix B.

This section discusses:

1. SABS, the program that digests the raw NFAC file and produces a variety of output files for processing by SPSS and other utility programs.

2. The SPSS programs currently written to provide statistical breakdowns on data files from SABS.

3. Projected additions to SABS and the SPSS programs.

4. Proposed distribution of results of statistical log processing.

3.1 SABS - ON-LINE CATALOG STATISTICAL FILE PROCESSING

SABS has been designed to accomodate a variety of processing requirements. The program contains numerous software 'switches' giving DLA analysts control over the processing logic. Switches also control the writing of output files, so that only the data required need be produced. Simple statements included in the SABS control file set the various switches and control the processing sequence for any particular run on SABS.

Two classes of procedures, or 'modes' may be evoked by statements in the control file, SCAN mode and NORM mode.

3.1.1 SCAN Mode

The only SCAN procedure currently available is the TIME SCAN. (Other SCAN procedures under development are described in section 3.3.)

The TIME SCAN procedure takes as input the raw (unsorted) NFAC file and optionally produces a histogram of system activity. The histogram is broken into intervals (specified in the procedure statement by the user). For each interval the histogram procedure outputs columns of characters representing the following values.

1. Number of active sessions.
2. Number of logons.
3. Number of logoffs.
4. Number of 'timeouts'.
5. Number of disconnects (without explicit logoff).
6. Number of commands issued by users.
7. Number of screens transmitted to users.
8. Amount of time the system spent in MWAIT (idle state).
9. Average response time for searches.
10. Number of calls to ADABAS.

Each interval also includes the time of day the interval ended, and the system run status at that time.

A file (NCDQ)[3] is generated simultaneously containing the same information as the histogram for further processing by the ACTSUM[4] SPSS program.

---

[3] All output files from SABS are described in Appendix C.

[4] All SPSS programs currently available are described in Appendix D.

3.1.2 NORM Mode

The NORM mode of operation for SABS uses the NFAC file sorted by the run (job) identifier, the user's session number, and the time the record was logged. SYNCSORT is used for this sorting, which essentially groups together all records of a particular user session in chronological order.

The current version of SABS permits only one sorted NFAC file to be input per run. SABS produces 13 output files during normal running (in addition to the 1 file produced by SCAN runs.)

Table 1 lists the file names and their general contents. A more complete description for each file may be found in Appendix C.

TABLE 1

Files currently generated by SABS

FILE NAME	CONTENTS	SPSS PROGRAM
NCDR	DISPLAY SUMMARY RECS	DISPSUM
NCDS	SESSION SUMMARY RECS	SESSUM
NCDT	SEARCH SUMMARY RECS	SRCHSUM
NCDU	COMMAND TRANS. RECS	COMMTRNS
NCDX	SCREEN TRANS. RECS	SCRNTRNS
NCDY	SUBSYSTEM TRANS. RECS	
NCDQ	ACTIVITY SUMMARY RECS	ACTSUM
NCDV	SEARCH TEXT RECORDS	
NCDW	ERROR TEXT RECORDS	
NCDZ	SELECTIVE RECORD FILE	
NFAO	HELP SUMMARY&TEXT	HELPREC
NFAP	ERROR SUMMARY RECS	ERRREC
NFAQ	BROWSE SUMMARY RECS	BRWSREC
NFAR	BROWSE TEXT RECORDS	

In NORM mode the SABS program reads one session at a time from the sorted NFAC file into an array. Each record of the session is evaluated in sequence and summary or transition records are produced for each event (i.e., search, display, etc.) using information from one or more monitor records.

As each record is completed it is written out to the appropriate file.

3.1.3 SABS Processing Sequence

The normal processing sequence for NFAC statistical files is as follows:

1. Processing is initiated by the monitor asynchronous output routines. (See Figure 2)

2. The sequential NFAC file is sorted by SYNCSORT into 'sessions', and the unsorted file is copied to tape for later reference.

3. SABS processing is started.

   a) The sequential NFAC file is run through SABS in SCAN mode. A histogram of overall system activity, as reflected in the file, is produced.

   b) The sorted NFAC file is run through SABS and the output files listed in Table 1 are generated.

4. The SABS output files are run through the appropriate SPSS program, resulting in a printed statistical report.

This entire processing sequence runs as a linked 'network' of jobs. Successful completion of one job of the network initiates the next job until processing is complete.

3.2 SPSS PROGRAMS

The names of the SPSS programs used to analyze output data from SABS are listed in Table 1 More complete descriptions of the current SPSS programs may be found in Appendix D.

Currently these SPSS programs do not do any sophisticated statistical analyses of the data from SABS. For the most part simple frequencies of nominal variables (i.e., code values) and descriptive statistics of interval variables (mean, mode, variance, max, min, standard deviation, etc.) are generated.

Currently, no recoding (or grouping of ranges of values) is done. Selection of records to be processed by particular SPSS procedures (based on variable values) is done only to differentiate between LOOKUP and COMMAND mode records in

most SPSS procedures. Some cross-tabulations between two variables are performed. (E.g., In transition records cross-tabulations are used to show the probabilities of moving from one state to another. See descriptions of COMMTRNS and SCRNTRNS in Appendix D.)

3.3 PROJECTED EXTENSIONS TO SABS

The SABS program is expected to continue growing and changing along with the changes to the on-line catalog. This section describes some of the additions to SABS currently under development.

3.3.1 ADABAS/NFAC Linking of Log Records

The major anticipated extension to SABS is a module that will allow linking of data between the NFAC monitor log file and the ADABAS log file. Variables to hold information from the ADABAS log have already been allocated to various SABS output files, and the design of the algorithms to make the 'many-to-one' match between records in two files is currently being designed.

3.3.2 Session Translation Scan

This extension to SABS SCAN mode would take as an input control card one to five session numbers and generate a printed listing of all the events in those sessions, translated from their NFAC coding into 'English'. Such a listing might look something like the following:

```
SESSION NUMBER 000012     TERMINAL ID = 051 B51 BK
LOGON TIME  08.35.21

EVENT                                                  SECONDS SINCE
                                                       LOGON

SCREEN SENT TO USER  #0002300                          000000.873

USER INPUT: 2                                          000005.432

(COMMAND MODE SELECTED)

SCREEN SENT TO USER  #0002330                          000005.981

USER INPUT: find su energy policy                      000032.123

RETRIEVAL REPORT:  356 ITEMS SYSTEMWIDE
                   RESPONSE TIME 000001.034

USER INPUT: help display                               000052.231

(HELP WORD FOUND IN GLOSSARY)

SCREEN SENT TO USER  #0001200                          000052.899

 USER INPUT:  display review
 ...
```

and so on.

3.3.3 On-Line Questionaires

When on-line questionnaire administration begins for the catalog the NFAC file will hold the responses. SABS is being expanded to select questionnaire responses and create output records in the same format as the records produced by processing the questionnaires used by the other project participants.

3.3.4 Comments Command

The comment command to be added to the on-line catalog will allow users to input free text comments that will then need to be selected by SABS and written out to a file for review by DLA Staff.

3.3.5 Concordances of Command Text

Several output files from SABS contain copies of user input command text. These text files will be processed through a concordance program to generate frequencies for individual keywords used for searching, browsing, etc. Ideally, the concordance will not be recreated for each run, but instead updated so that the concordance represents keyword utilization over an extended period of time.

Separate concordances could be generated for error text and valid command text, or for searches in specific indexes, etc.

3.4 DISTRIBUTION OF MELVYL STATISTICAL REPORTS

Reports on MELVYL use and performance are needed by DLA staff for many purposes. The following describes the schedule for running of SABS and the SPSS analysis programs, and for internal and external distribution of the resulting statistical reports.

3.4.1 Daily Runs

As described in section 2, SABS processing is initiated automatically whenever one of the NFAC direct output files becomes full. This will occur about once per day. The control file set up for normal running of SABS includes the following steps:

1. SABS is run in SCAN mode using the unsorted NFAC sequential file. A histogram of system activity with a 300 second (5 minute) interval and the NCDQ activity summary file are produced.

2. On completion of the SCAN run, SABS switches to NORM mode and processes the sorted NFAC file. All output files are produced. (See Table 1)

Upon completion of SABS, the appropriate SPSS programs are initiated and printed reports are produced.

A program is run to print copies of search commands, browse commands, command errors, and comments for review by DLA staff.

The SABS output files are then copied to weekly cumulative files.

3.4.2 Weekly Runs

SPSS runs on the weekly cumulative files will be done every monday morning, following the cumulation of any Friday and weekend NFAC files. These runs will carry out more complete analysis of the data for the particular week. The data from the weekly file will be copied to cumulative tape files for later reference.

3.4.3 Monthly Runs

On the first of each month the cumulated tape files will be processed to provide an overview of the month's activity.

3.4.4 Internal Report Distribution

Output from the various processing steps described above will be available to DLA staff both on-line and as a binder of printouts from the analysis programs.

The binder will include the following items:

1. A description of binder contents and instructions on how to access files on-line.
2. Histograms of daily system activity (retained for one week in the binder then removed for storage elsewhere.)
3. Printouts of daily text files (including search, browse, and unusual help commands, as well as text of command errors.)
4. Printouts of daily SPSS runs.
5. Printouts of the most recent weekly SPSS runs.
6. Printouts of the most recent monthly SPSS runs.

3.4.5 External Report Distribution

During Phase II of this project analyses of information collected on catalog use and users will be distributed to project participants.

Requests for raw monitor data or SABS output file data will be considered on a case-by-case basis.

Section 4

ON-LINE QUESTIONNAIRE ADMINISTRATION

Early in the planning for this project it was decided that a single instrument would be developed to facilitate comparisons between the user and non-user populations of the 15 participating library systems.

The uniform instrument was developed jointly by RLG, OCLC, Joseph Matthews Associates, and DLA, coordinated by Doug Ferguson of RLG. The final form of the questions agreed upon in a meeting by all participants in Chicago on July 15, 1981. The questionnaire and the development process are more fully described in the RLG Final report on Phase I (Still in preparation).

In order to reduce the costs of questionnaire administration and at the same time expand the respondent base, DLA decided to implement the uniform instrument as an on-line subsystem of the catalog.

Early planning for the DLA on-line interaction monitoring included provisions for on-line administration of questionnaires. (The software being developed for on-line questionnaire administration was originally planned for internal studies of on-line catalog use and/or other library survey research. Since the questions asked may be easily changed to accomodate a variety of study goals, the catalog's questionnaire subsystem will probably see more use as the needs for information about UC library users change over time.)

The questionnaire administration software is being developed to facilitate selective sampling by DLA analysts. The sampling capabilities will include:

1. An 'On-Off' software switch to enable or disable the questionnaire subsystem while the catalog is running.

2. The ability to select a particular set of terminals for study. (E.g. only dial-up users, only users at a particular campus, only terminals in branch libraries, etc.)

3. The ability to administer questionnaires only to particular groups of users. (e.g., only students, faculty, library staff, etc.)

The questionnaire administration software will incorporate several 'decision points' to facilitate sampling. The logical flow of the questionnaire subsystem is illustrated in Figure 3

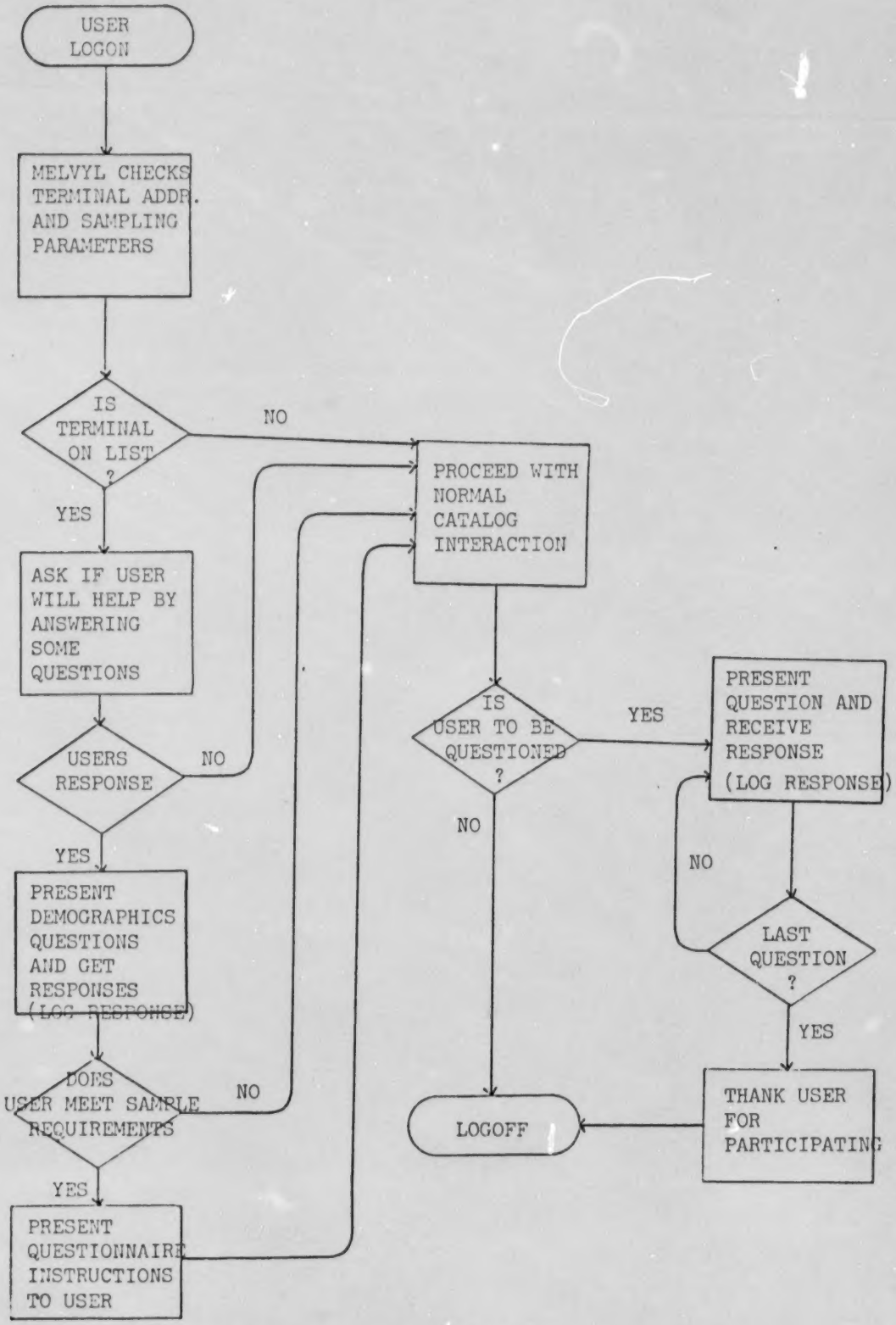
Initially, the questionnnaire subsystem will interact with users only at the beginning and end of their catalog session. However, the questionnaire subsystem will eventually be expanded to allow administration of questions at a number of well defined points during a session. For example, questions about the readability of a record display format could be asked immediately following the actual display, perhaps on the same screen.

The logic incorporated into the questionnaire subsystem will guide respondents through a set of related questions, eliminating questions logically excluded by their responses. (E.g. If users indicate that there is no local printer attached to their terminal, no other questions about local printing will be asked of them.)

The on-line questionnaire subsystem will immediately check user responses for validity, and reprompt the user when the response is out of the acceptable range of values, contains invalid characters, or has other errors.

All responses to the questionnaires will be logged along with the rest of the user's interaction in the statistical data log NFAC. During batch processing of the NFAC file questionnaire responses will be accumulated in records in a format identical to the records produced during automated processing of the paper questionnaire forms used by the other project participants. Thus, the data generated by all project participants can be integrated into a single data base.

Figure 3. Questionnaire Subsystem Logic Flow



Section 5

MELVYL PRETEST: SOME EARLY RESULTS

A 'Pre-Test' version of the Prototype On-Line Catalog has been available for several hours each weekday during the months of May, June, and July 1981. The primary purpose of the pre-test has been to familiarize campus librarians with the on-line catalog and to get their comments and suggestions, before the catalog is released to the general public. However, at several locations terminals were made available to all library users. After some initial testing and modifications, the monitoring and analysis programs were run daily beginning June 1, 1981. This section describes some early results from analysis of the data collected. It must be stressed that these observations are not to be considered 'final' in any sense. Two major factors affecting the data must be considered:

1. The 'novelty factor.' Having the library catalog on-line is a new experience, and people have a tendency to 'play' with new electronic devices. Observation revealed that many users were obviously looking up all their favorite authors.

2. The population is predominantly librarians, although many 'normal' library users are included.

5.1 THE AVERAGE USER SESSION

Based on an analysis of 6911 user sessions collected from June 1 to July 7, 1981, the following section presents the 'average user session'.

54% of the users chose 'LOOKUP' mode for their sessions, and the remainder chose command mode.

The mean length of time that a user spends at the terminal (from logon to logoff) is 14 minutes, 19 seconds (859.018 seconds). This mean is skewed by a few very long sessions. 71.2% of all sessions last less than 10 minutes and only 4.4% of all sessions last longer than 1 hour.

The mean number of commands (of all types, i.e., search, display, and help commands, also responses to prompts, car-

riage returns to continue displays, etc.) issued in a session is 25.707.

The user issues a mean of 5.91 searches per session. (The maximum observed was 198 searches in one session. Probably an example of the novelty factor.)

Only 24% of the sessions used the help facilities. (The Help subsystem was still being developed during the pre-test period, and had not been publicized to any great extent.) Those who did use Help issued a mean of 1.5 unqualified 'HELP' requests and 3.1 Help requests with a glossary term (e.g. 'HELP INDEXES'.)

Only 3.5% of the users made the same error three times in a row and were automatically sent a Help screen.

A mean of 11.76 display screens were sent to users per session.

A rather surprizing 64% of the users made no errors at all during their sessions (This includes command syntax errors, logical errors --such as trying to display when no records have been retrieved --, and issuing unrecognizable commands.) The other 36% made an average of 3.6 errors during their sessions.

Only 13% of the users made use of the BROWSE command, but those who did used it an average of 3.76 times.

5.2 DISPLAYS

During the pre-test period 48,798 records of displays were collected, of these 64.8% were from command mode sessions and 35.2% were from 'LOOKUP' mode sessions.

For the LOOKUP mode sessions, 42.4% of the displays were in review format (the automatic default), 35% were in 'Long' format, and 21.4% were in 'Brief' format. In 57.5% of the displays, the user took 25 seconds or less to absorb the information on the screen and issue the next command. 88.7% took less than 1 minute.

For users who chose command mode, 53.4% of their displays were in 'Brief' format (the default), 17.1% were in 'Long' format, and 6.8% were in 'Review' format. Of the remainder, 14.4% of the display screens sent to command mode users were displays of lists of headings retrieved by 'BROWSE' commands, 4.9% were displays of individual fields or groups of fields from the MARC record, and 3.5% were full MARC displays.

For 54.6% of the command mode displays it took the user less than 15 seconds to absorb the information in the display and issue a new command. In 92.4% of the cases 'think time' was less than 1 minute. (There were several cases where the user obviously just walked away from the terminal after a display without logging off, where 'think time' was more than 20 minutes.)

5.3 SEARCHING PATTERNS IN THE ON-LINE CATALOG

Perhaps the most interesting and significant observations deriving from the pre-test are the ways people search for things in on-line catalogs.

There have been an number of studies of the way people use card catalogs in libraries, but studies of the on-line catalog are just beginning. The early results presented here indicate that on-line catalogs are being used in a qualitatively different fashion from card catalogs.

In a major study at Yale University[5] it was found that 73% of card catalog users were searching for 'known items' and only 16% were conducting subject searches. Lipetz suggested that many of the so-called known item searches were actually disguised subject searches, making the 'real' breakdown 56% known item and 33% subject searches. A study by Palmer[6] at the University of Michigan found a similar 70% known item searches.

Swanson,[7] in reviewing 16 catalog use studies (representing approximately 8000 searches) suggests that the average searching pattern is 60% known item and 38% subject searching.

These studies have obvious limitations, they cannot easily directly observe what the user is doing, and must rely on interviews or questionnaires to have the users describe their intentions in searching the catalog. Through the use of monitoring we are able to directly observe what indexes the users are searching.

---

[5] Ben-Ami Lipetz, 'Catalog Use in a Large Research Library': 1972.

[6] R. P. Palmer, 'Computerizing the Card Catalog in the University Library': 1972.

[7] D. R. Swanson, 'Requirements Study for Future Catalogs' : 1972

The searching patterns of users of the on-line catalog are rather dramatically different from those observed in the card catalog studies.

For 11,963 command mode searches using only a single index, 45.2% were subject searches, 23.2% were personal author searches, 3.9% were corporate author searches, and 21.9% were title searches (which may or may not be 'disguised' subject searches looking for keywords in titles). The remainder were searches on series titles or uniform titles, which again, may be disguised subject searches.

For 1098 command mode searches using 2 indexes combined by Boolean operators, 41.4% used a combination of subject and title and 41.0% used a combination of personal or corporate author and title (definitely known item). Of the remainder, 8.1% used a combination of author and subject indexes and 5.9% used combinations of personal and corporate authors.

For 6791 searches in 'LOOKUP' mode, 49.1% were subject searches, 37.9% were author or author/title searches (known item), and 12% were title searches alone.

These early observations require further testing with a more diverse user population (to be carried out in Phase II of this project), and the supplementary information from on-line questionnaires that will allow breakdowns of the data by user characteristics.

5.4 OBSERVATIONS AND CONCLUSIONS

LOOKUP mode seems be be fulfilling its function as an easy method of access for users who wish to 'find a book' quickly. LOOKUP mode users make one or two searches (of the 23,719 searches observed, 70.6% were issued in command mode, and only 29.4% were issued in LOOKUP mode, yet 53% of all sessions were LOOKUP mode sessions.)

The differences between searching patterns in the on-line catalog and the card catalog suggests (what many librarians have suspected) that the lack of subject searching in card catalogs is not because users don't want to do subject searches, but because the way subjects are presented and arranged in card catalogs is not conducive to use. (As Mooer's law of information retrieval states, "An information retrieval system will tend not to be used when it is more painful and troublesome for a customer to have information than for him not to have it."[8]) When users are given a 'better method' to get at the subjects they are interested in, they will tend to do subject searches, and the on-line catalog can provide a less 'painful and troublesome' method.

The obvious conclusion to be drawn from these early observations is that the on-line catalog is not going to be used just like the card catalog, and on-line catalog designers should avoid relying entirely on studies of card catalogs for indications of how to proceed. We should not be re-creating the card catalog on-line, we should be creating a new, more sophisticated and at the same time less 'painful' way of exploring the riches of our libraries.

---

[8] C. N. Mooers, 'Mooers Law or, Why Some Retrieval Systems Are Used and Others Are Not.' American Documentation vol. 11, no. 3 (1960) p. ii.

BIBLIOGRAPHY

Lipetz, Ben-Ami. 'Catalog Use in a Large Research Library.' Library Quarterly. vol. 42 (1972) p.121-139.

Palmer, R. P. Computerizing the Card Catalog in the University Library: A Survey of User Requirements. Littleton, Colo.: Libraries Unlimited, Inc., 1972.

Penniman, W. D. and Dominick, W. D. 'Monitoring and Evaluation of On-Line Information System Usage.' Information Processing and Management. vol. 16 (1980) p. 17-35.

Swanson, D. R. 'Requirements Study for Future Catalogs.' Library Quarterly. vol. 42 (1972) p 305-315.

Appendix A

THE MELVYL PATRON INTERFACE

The UC Prototype On-Line Catalog provides public access to the library resources of the nine campuses of the University of California using the UC Union Catalog Data Base (UCDB), which contains over 730,000 records of works cataloged by UC campuses through automated technical processing systems (e.g., RLIN and OCLC).

A.1 THE PATRON INTERFACE

At the heart of the catalog is the patron interface system, consisting of the programs that mediate all interaction between the user, the computer, and the data base.

The patron interface will support user interaction at three different levels, or 'modes.'

A.1.1 Look-Up Mode

At the most basic level, Look-Up mode provides a very simple access method for inexperienced users. It combines choices from menus and computer-prompted 'fill-in-the-blank' questions to access the author, title and/or subject indexes. All responses to the user include simple explanations of the options available and instruction on how to get Help, or further instructions.

A.1.2 Tutorial Mode

Tutorial mode (still under development) will introduce the user to the command language level of system interaction by step-by-step explanations of how (and why) commands are constructed. Tutorials will be available on each of the indexes, on the display formats, on how to create bibliographies from retrieved records, etc.

A.1.3 Command Mode

Command mode is a concise query and data manipulation language for experienced users. It gives the user the ability to conduct very sophisticated searches and to manipulate retrieved records with a minimum of computer prompting.

These three modes of interaction with the system are more fully described in Division of Library Automation Working Paper No. 7, University of California Prototype On-Line Catalog: Preliminary Specifications for the Patron Interface (July, 1980).

A.2 ACCESS POINTS

The patron will be able to search for catalog records through several different categories of access points.

A.2.1 Primary keyword search indexes

Most searching in the catalog will be done through the primary keyword indexes. Keyword searches may be performed on single words or Boolean combinations of words. Those that can be used to initiate a search are known as primary keyword searches, and the indexes include:

1. Personal Author, including main and added entries.
2. Corporate/Conference Author, including main and added entries.
3. Title, including all title fields, all series statements unless traced differently, and all series tracings.
4. Uniform Title, including all uniform title fields.
5. Series, including all series statements unless traced differently, and all series tracings.
6. Subject, including all subject tracings.

Some early results on the frequency of use of these access points may be found in section 5 of this report.

A.2.2 Qualifying (secondary) search indexes

A second category of search index contains items which are not unique enough to be used efficiently to access the catalog directly, but which are useful in refining a set of retrieved records. Qualifying (or secondary) search indexes may be used only after a set of records has been retrieved using one or more primary search indexes. Currently these qualifying search indexes include:

1. Publication date
2. Publisher
3. Place of publication

Eventually, the following secondary indexes will be available:

1. Edition
2. Language of the text of the work
3. Chronological coverage of the work
4. Geographic area covered by the work

A.2.3 Fixed format control numbers

The simplest type of search will be through fixed format control numbers. For these searches the patron will enter the search index and the number desired. In most cases, each number will retrieve one record. These 'number indexes' are primarily for the benefit of librarians searching for known items. The control numbers available for access to the catalog are:

1. Union Catalog record number (URECNO)
2. LCCN (Library of Congress Catalog Card Number)
3. ISBN (International Standard Book Number)
4. OCLC number (record number in the OCLC technical processing system)
5. RLIN number (record number in the RLIN technical processing system)
6. ISSN (International Standard Serial Number)

A.3 SEARCHING IN THE PROTOTYPE ON-LINE CATALOG

The UC Prototype On-Line Catalog will support several forms of searching. Every search will begin with the verb 'FIND', although in Look-Up mode the verb is supplied by the system rather than the user. Following the verb will be the search 'arguments' consisting of the name of the index to be searched, and the value to be located. The index names may be abbreviated. (For example, 'FIND TI LIBRARY' would locate all books with the keyword 'library' in their titles.)

An initiating search (one beginning with 'FIND') may be continued after retrieval has taken place by starting the next search command with one of the Boolean operators (AND, OR, or NOT). This modifying search will be conducted by the system as if the initial search and modifying search were issued as a single command (i.e., only the records that meet the criteria of both search statements will be retrieved). The effects of a modifying search may be negated by issuing a 'BACKUP' command, which will return the search results that were in effect before the latest modifying search was started.

A.3.1 Boolean searching

Multiple search arguments may be combined in the same search by use of Boolean logical operators to interrelate arguments. The Boolean operations supported by the prototype on-line system are:

1. AND. To satisfy a search with two arguments connected by a Boolean 'AND,' an item must satisfy both arguments. (E.g., to satisfy the search 'FIND PA DICKENS AND TI TWIST', only books with author's names including the word 'DICKENS' which also have titles including the word 'TWIST' would be retrieved.)

2. OR. To satisfy a search with two arguments connected by a Boolean 'OR,' an item must satisfy either one or the other (or both) of the arguments. (E.g., to satisfy the search 'FIND PA DICKENS OR PA BALZAC', books having author's names including either the word 'DICKENS' or the word 'BALZAC' (or both) would be retrieved.)

3. NOT. To satisfy a search with two arguments connected by Boolean 'NOT,' an item must satisfy the first argument and fail to satisfy the second. (E.g., to satisfy the search 'FIND PA DICKENS NOT TI TWIST', all books having authors' names including the word 'DICKENS' would be retrieved, except for those that also contained the word 'TWIST' in the title.)

When two or more keywords are specified in a search argument without an explicit Boolean operator, 'AND' is supplied by the system.

A.3.2 Keyword modifications

Any search on a keyword may be refined or broadened to make the search more or less specific. The following modifications will be allowed:

1. Truncation. This provides a generic search for all keywords that begin with the characters entered. A title search on the truncated keyword 'libr' will retrieve all titles containing any word having 'libr' as its first four letters. There will be a lower limit to the number of letters that can be used in a truncation search.

2. Phrase. A search on a phrase will return only records containing the keywords in the order entered and in adjacent positions. A title phrase search on 'air pollution' will retrieve the title 'Air Pollution in California' but not 'The Pollution of California's Air.'

3. Literal search. A literal search will retrieve only records that exactly match the entered search key. A literal title search on 'Sleeping Beauty' will retrieve only titles that consist wholly of the words 'Sleeping Beauty.'

A.4 DISPLAY FORMATS

Catalog records may be displayed in a number of different formats, which can be selected by the user.

A.4.1 Review display

This format is useful for quickly scanning a large number of retrieved records. It provides the main entry, title (truncated if necessary) and the publication date on one or two lines of the terminal screen. Each record displayed is numbered so that the user can select one or more for expansion. The records are alphabetized by the main entry, title and date. (A feature to be added to the catalog will permit users to specify their own sort order for displays.)

A.4.2 Brief display

A brief display consists of the main entry, full title, place, publisher, date, and call numbers for each book retrieved. Each record is numbered so that the user can select one or more for fuller display. This format would be requested by the user for detailed, but quick, examination of records when the review display was not precise enough, or when the call number is needed to find books on the shelf.

A.4.3 Long display

The long display format presents complete bibliographic information on each union catalog record with 'labels' for each field in the record (i.e., 'Author', 'Title', 'Notes', 'Subjects', etc.).

A.4.4 MARC display

The MARC format provides a display of a single union catalog record in the full MARC format, with leader and all fields, tags, and indicators, similar to the microfiche union catalog register display, including all local (campus) variant headings.

A.4.5 Tags display

This display format allows the user to look at one or more MARC fields, specified in the display command. For users not familiar with the MARC format, several key fields may be displayed by specifying their contents. (e.g., 'DISPLAY AUTHORS' will cause a display of all the 1xx and 7xx fields, 'DISPLAY SUBJECTS' will show all 5xx fields.)

A.5 BROWSING

The Browse capability of the catalog gives users the ability to look at a complete listing of the 'headings' in the catalog that contain a particular word. For example, the command 'BROWSE SU ENERGY' would retrieve a numbered list of all subject headings in the catalog that contained the word 'ENERGY' regardless of where in the heading the word occurred. The user can choose which of the browsed headings express his interests best and command the system to 'SELECT' one or more. MELVYL then conducts a search for all the books in the database that are indexed by the specified headings.

A.6 HELP

At any time during a session, the user may ask for help. If the user just types 'HELP' the system will report what has just occurred in the session, and suggest how to continue or get more help. If the user types 'HELP' followed by any of the command words (or several synonyms) the system will provide a full explanation of the command and how to use it. If a user makes the same error in interacting with the system three times consecutively, the system assumes that help is needed and will automatically send out an appropriate explanation, just as if the user had typed 'HELP'.

Appendix B

USER TRANSACTION LOG: CODES AND CONTENTS

The NFAC user transaction log file contains records of varying length (up to 254 characters). Each record contains a fixed length portion providing identification of the user session associated with the record, some time-stamps, and other information necessary to processing of the records.

This appendix describes the contents of the fixed portion of the NFAC records and the variable portions generated by each different subsystem of the catalog, along with the codes differentiating between record types.

B.1 VARIABLES COMMON TO ALL RECORDS

The following table lists the variables that make up the fixed portion of each NFAC record, followed by short descriptions of the contents of the variables.

VARIABLE NAME	LENGTH IN BYTES	TYPE
PLILEN	2	BINARY
JOBID	6	CHARACTER
DATE	6	NUMERIC CHARACTER
TIMEOD	6	NUMERIC CHARACTER
TIMEHLD	4	BINARY
TIME_LASTIN	4	BINARY
TERMID	3	NUMERIC CHARACTER
TERMADD	3	CHARACTER
SESSION	6	NUMERIC CHARACTER
FILLER	9	CHARACTER
STRUC_ID	1	CHARACTER
MSGCODE	5	CHARACTER
TEXT	195	CHARACTER (VARYING)

1. PLILEN : The length of the record (fixed + variable portion.)

2. JOBID : An identifier for a particular run of the catalog.

3. DATE : The date the record was written.

4. TIMEOD : Approximate time of day the record was written.

5. TIMEHLD : A time stamp from the internal system clock, recorded when the record is written out. (This variable is in 1024 units per second of clock time, a subroutine in the SABS program can convert this time stamp to clock time with accuracies up to +- 1 second, and compute elapsed times between two time stamps accurate up to +-.001 seconds.)

6. TIME_LASTIN : A time stamp (in the same form as TIMEHLD) that is recorded when the user enters a command. (For each NFAC record, TIME_LASTIN gives the time the last command was entered for the session.)

7. TERMID : The logical terminal identification number assigned by the catalog system at logon.

8. TERMADD : The telecommunications port address for the terminal (in hexadecimal.) For 'hard-wired' terminals this address may be 'decoded' to indicate the actual location of the terminal.

9. SESSION : The session number assigned by the catalog system at logon.

10. FILLER : Space for record expansion.

11. STRUC_ID : An identifier indicating the PL/I data structure used to format the TEXT portion of the record.

12. MSGCODE : An alphanumeric code indication the origin and contents of the text portion of the record. (These codes are described below.)

13. TEXT : The variable portion of the NFAC record used by the different subsystems to include a variety of data.

Each major subsystem of the on-line catalog produces different types of NFAC records. The following sections describe the records generated by each subsystem, their associated MSGCODES and contents.

B.2 QCONTRL SUBSYSTEM RECORDS

The format for NFAC MSGCODEs generated by the QCONTRL subsystem is 'COnnn', (where 'nnn' is a three digit number.) The MSGCODEs may be grouped into general classes based on their numbers.

C0001-C0199 include records related to user input and non-error conditions encountered during normal processing of commands.

C0200-C0399 include records related to user input errors and exceptional conditions raised during normal processing of commands.

C0500-C0599 include record on the 'housekeeping' by the controller, and periodic reports of cumulative variables maintained by the controller.

Most of the possible MSGCODE values in these ranges are not currently used, those that are are described below.

1. C0001 : Records with this code contain a copy of the user's command, and it's length (number of characters) in the variable portion of the record.

2. C0003 : This code indicates that a terminal activation request was received by the system (logon or START of a particular session.) Terminal location and identification information is included in the variable portion of the record.

3. C0004 : This code indicates a terminal deactivation by explicit user logoff.

4. C0005 : This code indicates a terminal deactivation by time-out. (Although not presently in effect, the catalog system has the capability of 'logging off' any terminal which has not shown activity for a set amount of time.)

5. C0006 : This code indicates terminal deactivation by disconnection. (i.e., a dial-up user just hangs up the phone without logging off, power to a terminal was cut off, etc.)

6. C0100 : This code indicates that a screen was transmitted to the user. (For LOOKUP, HELP, questionnaire, and error message screens. Does not include display screens.) The screen number is included in the variable portion of the record.)

7. C0201 : This code indicates a terminal read error. (That is, the system received garbage from a terminal.) Usually records with this code are caused by telecommunications line 'noise.' May be used of identify bad lines or terminals.

8. C0202 : This code indicates that a user was interrupted during input by a system broadcast message of some sort.

9. C0203 : This code indicates the use of illegal tabbing (instead of spacing) by the user.

10. C0204 : This code indicates an excessive number of 'breaks' being sent from a particular terminal.

11. C0206 : This code indicates that a terminal had to be denied logon because the system is operating at full capacity.

12. C0207 : This code indicates that the terminal output buffer space was full and is being compressed.

13. C0208 : This code indicates that terminals were idled because storage limits were reached.

14. C0209 : Terminals attempting to logon were automatically logged off because ADABAS was not running.

15. C0500 : This code indicates a 'housekeeping record' that includes the following information:

    a) Number of active terminals on system.

    b) Number of breaks received since last record was written.

    c) Number of ADABAS calls issued since last record.

    d) Amount of time spent in MWAIT (an idle state entered by the catalog when there is no activity from users) since last record was written. (in 1024 units per second.)

B.3 QDISPLA SUBSYSTEM RECORDS

All records generated by the QDISPLA subsystem have the MSGCODE DI000. A record is generated for each display screen sent to the user. The variable portion of the NFAC record contains the following information:

1. STAT CURRENT RANGE: When multiple ranges of records are specified in a command (e.g., DISPLAY 1-5,7,21-30) this variable indicates which of the range groups is currently being displayed.

2. STAT CURRENT REC: Sequence number of the record currently displayed. (If there is more than one record per screen, the sequence number of the first record of the screen is logged.)

3. STAT SCREEN# : Sequence number of the screen of records currently displayed. (i.e. first screen = 1, second screen = 2, etc.)

4. STAT REC CNT : The number of records displayed on the current screen.

5. STAT BREAK BIT : A bit flag indicating that the user hit the break key during the display (before the display was completed.)

6. STAT FIRST REC : A bit flag indicating that the screen being displayed contains the first record of the display request.

7. STAT LAST SCREEN : A bit flag indicating that the screen being displayed is the last screen of the display request.

8. STAT PARTIAL REC : '1' if a partial record is displayed, otherwise '0'.

9. STAT CURRENT FMT : A code indicating the display format used in the display. (i.e., brief, long, review, MARC, tags, or labeled tags display)

B.4 QHELP SUBSYSTEM RECORDS

The QHELP subsystem handles all errors and error messages for the on-line catalog, as well as HELP request processing. The MSGCODE format for QHELP generated NFAC records is 'HEnnn'. The following codes are currently used:

1. HE001 : Help responded to an unqualified help request. A code indicating the user's status at the time of the help request is recorded in the variable portion of the record. These codes indicate such conditions as 'search completed with no records retrieved', 'last search request had syntax errors', etc.)

2. HE002 : Help responded to a 'HELP term' request, and the term was found in the on-line glossary. A code indicating the user's session status and the term requested are logged in the variable portion of the record.

3. HE003 : Help responded to a 'HELP term' request, and the term was not found in the on-line glossary. A code indicating the user's session status and the term requested are logged in the variable portion of the record.

4. HE004 : Help sent out a brief error message when a user error was detected. The user's session status (an error code in this case) and a copy of the command causing the error are logged in the variable portion of the record.

5. HE005 : Help sent out a help screen because 3 identical consecutive errors were encountered. The user's session status (an error code) and the command causing the error are logged in the variable portion of the record.

B.5 QLOOKUP SUBSYSTEM RECORDS

The QLOOKUP subsystem acts as a front-end to the on-line catalog for inexperienced users. Through the user's responses to menu and 'fill-in-the-blank' screens QLOOKUP constructs commands that are then processed as if they came directly from the user. The MSGCODE format of QLOOKUP records is 'LKnnn'.

1. LK001 : This code indicates a search command was constructed by QLOOKUP. the command mode search request text is logged in the variable portion of the record.

2. LK002 : This code indicates a display command was constructed by QLOOKUP. The command mode display request text is logged in the variable portion of the record.

3. LK003 : This code indicates that a user made an incorrect response to a LOOKUP menu selection prompt.

B.6 QSEARCH SUBSYSTEM RECORDS

The QSEARCH subsystem parses incoming search commands and formulates the ADABAS calls required to complete those searches. The QSEARCH MSGCODE format is 'SEnnn'.

1. SE001 : This code indicates that the user issued an initiating search statement. (e.g. 'FIND PA DICKENS')

2. SE002 : This code indicates that the user issued a secondary or modifying search using primary search indexes. (e.g. 'AND TI OLIVER TWIST')

3. SE003 : This code indicates that the user issued a modifying search using secondary indexes. (e.g. 'AND DATE 1976-')

4. SE004 : This code indicates that the user issued a location processing command to limit a retrieved set of records to those held at a particular location or type of library. (e.g. 'AT BERKELEY', 'AT MEDICAL', etc.)

5. SE100 : This code indicates a retrieval report. The variable portion of the record includes the following data:

   a) Number of records retrieved (overall).

   b) Location code if search was limited by location.

   c) Number of records at the specified location.

6. SE200 : This code indicates the search is a 'long search' requiring special processing.

7. SE201 : This code indicates that the user issued a 'BACKUP' command to restore the results of an initial search changed by a modifying search.

8. SE300 : Indicates that the search originated as a 'SELECT' command, where the user chose one or more headings displayed in a BROWSE listing for retrieval.

Appendix C

SABS OUTPUT FILES

This appendix describes the files generated by SABS. Each file description includes a table giving the format of the variables and their placement in the record (as they appear to SPSS for SPSS input records) and a section describing the contents of each variable.

C.1 NCDQ - ACTIVITY SUMMARY FILE

VARIABLE	FORMAT	COLUMNS
YEAR	F 2. 0	1- 2
MONTH	F 2. 0	3- 4
DAY	F 2. 0	5- 6
HOUR	F 2. 0	7- 8
MINUTE	F 2. 0	9- 10
SECOND	F 2. 0	11- 12
SYSTAT	A 1	13- 13
IDLE	F10. 3	14- 23
TERMS	F 4. 0	24- 27
ADACALLS	F 4. 0	28- 31
AVGRSP	F10. 3	32- 41
SCREENS	F 3. 0	42- 44
COMMS	F 3. 0	45- 47
LOGONT	F 4. 0	48- 51
LOGOFFT	F 4. 0	58- 61
TIMEOUTT	F 4. 0	68- 71
DISCONT	F 4. 0	78- 81

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. HOUR, MINUTE, and SECOND: These three variables contain the time of day that the NFAC record providing a basis for this record was generated. The variables are included to allow aggregation of records by time of day in SPSS processing.

3. SYSTAT: System status at the time record was generated, usually either N for normal, T for terminal (last record of a run), or S for start (first record of a run).

4. IDLE: The amount of time in seconds and fractions of a second that the system spent in MWAIT idle state during a particular interval.

5. TERMS: Number of terminals (sessions) active at the end of an interval.

6. ADACALLS: Number of calls made to ADABAS during the interval.

7. AVGRSP: Average response time for completed searches made during the interval.

8. SCREENS: Number of screens transmitted to users during interval.

9. COMMS: Number of commands (of all types) issued by users during interval.

10. LOGONT: Total number of logons during the interval (the record may be expanded to include number of logons by location, i.e., Berkeley hardwire, dial-up, etc.).

11. LOGOFFT: Total number of logoffs (or normally ending sessions) during the interval.

12. TIMEOUTT: Total number of terminals timed out during interval (timeout is not currently enabled, so this variable is always null).

13. DISCONT: Total number of terminals that were disconnected from the system without explicit logoff command (i.e., a dial-up just hanging up the phone without logging off).

C.2 NCDR - DISPLAY SUMMARY FILE

VARIABLE	FORMAT	COLUMNS
YEAR	A 2	1- 2
MONTH	A 2	3- 4
DAY	A 2	5- 6
SESSION	F 6. 0	7- 12
LOGNO	F 4. 0	13- 16
USERTYPE	A 2	17- 18
SRCHMOD	A 1	19- 19

FORMAT1	A 1	20-	20
FORMAT2	A 1	21-	21
FORMAT3	A 1	22-	22
NFORM	F 7. 0	23-	29
SORTTYPE	A 1	30-	30
SORT1	A 2	31-	32
SORT2	A 2	33-	34
SORT3	A 2	35-	36
DISPTIME	F10. 3	37-	46
USERTIME	F10. 3	47-	56
ADACPU	F 9. 2	57-	65
ADATOTX	F 4. 0	66-	69
ADADATX	F 3. 0	70-	72
ADAASSX	F 3. 0	73-	75
ADAWRKX	F 3. 0	76-	78
RANGE	F 2. 0	79-	80
ERRS	F 4. 0	81-	84
CURREC	F 3. 0	85-	87
RECCNT	F 2. 0	88-	89
SCREEN	F 2. 0	90-	91
PART	A 1	92-	92
TERMIN	A 1	93-	93

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. SESSION: This is the number assigned by the catalog to a particular user at logon.

3. LOGNO: This is the sequence number within a particular session of the NFAC record providing the basis for each record of this file. This variable is included to aid in linking records from different files that relate to the same event.

4. USERTYPE: USERTYPE currently is not filled in by SABS. It is intended to contain a code indicating the user's status (e.g., student, faculty, librarian, etc.) when that information becomes available through on-line questionnaires.

5. SRCHMOD: This variable contains a code to indicate the searching mode set for a particular user at the time the record was generated. (E.g., LOOKUP, COMMAND, or not yet set.)

6. FORMAT1: Format set by default or by user command for a particular session.

7. FORMAT2: A code indicating when the format for a display was specified or by default (i.e., just 'display' was used.

8. FORMAT3: Display format used in the display.

9. NFORM: Number of records formatted (or to be formatted), equal to the last retrieval for the session.

10. SORTTYPE: Whether display sort order was by default, or specified (by an 'ARRANGE' command.)

11. SORT1 to SORT3: Sort keys used (up to three) in the order specified (or the default order: Main Entry, Title, Imprint Date.

12. DISPTIME: Response time in seconds and fractions of a second from the time the user pressed return at the end of the display command to the time the display record was logged (immediately before the display is transmitted).

13. USERTIME: The elapsed time between the time the NFAC display record was logged and the next command is input by the user (i.e., this variable includes the time it takes for a user to read and absorb, and respond to a display).

14. ADACPU: ADABAS CPU time consumed by the display command processing (not currently filled in).

15. ADATOTX: Total ADABAS EXCPs used in execution of the display command (not currently filled in).

16. ADADATX: ADABAS data storage EXCPs used in execution of the display command (not currently filled in).

17. ADAASSX: ADABAS associator EXCPs used in execution of the display command (not currently filled in).

18. ADAWRKX: ADABAS work storage EXCPs used in execution of the display command (not currently filled in).

19. RANGE: When multiple ranges of records are specified in a command (e.g., DISPLAY 1-5,7,21-30) this variable indicates which of the range groups is currently being displayed.

20. ERRS: All error reporting has been allocated to records from QHELP, so this variable will not be filled in.

21. CURREC: Sequence number of the record currently displayed. (If more than one record per screen, number of the first record of the screen is logged.)

22. SCREEN: Sequence number of the screen of records currently displayed (i.e., first screen = 1, second screen = 2, etc.).

23. PART: '1' if a partial record is displayed, otherwise '0'.

24. TERMIN: This variable contains a code to identify the way the display was ended. (I.e., 'C': user entered a new command, 'N': last record of the retrieved set or last record of display request set was displayed, 'Z': system crashed, 'B': user hit break during the display.)

C.3 NCDS - SESSION SUMMARY FILE

VARIABLE	FORMAT	COLUMNS
YEAR	A 2	1- 2
MONTH	A 2	3- 4
DAY	A 2	5- 6
SESSION	F 6. 0	7- 12
TERMLOC	A 3	13- 15
USERTYPE	A 2	16- 17
SRCHMODE	A 1	18- 18
STARTHR	F 2. 0	19- 20
STARTMIN	F 2. 0	21- 22
STARTSEC	F 2. 0	23- 24
ENDHR	F 2. 0	25- 26
ENDMIN	F 2. 0	27- 28
ENDSEC	F 2. 0	29- 30
ENDTYPE	A 1	31- 31
LENGTH	F10. 3	32- 41
ADACPU	F 9. 2	42- 50
ADATOTX	F 4. 0	51- 54
ADADATX	F 3. 0	55- 57
ADAASSX	F 3. 0	58- 60
ADAWRKX	F 3. 0	61- 63
COMMANDS	F 3. 0	64- 66
BREAKS	F 3. 0	67- 69
PARMTOT	F 3. 0	70- 72
PARMTRM	A 3	73- 75
PARMFRM	A 1	76- 76
PARMLIB	A 1	77- 77
PARMLAN	A 3	78- 80
HELPTOT	F 3. 0	81- 83
HELPUNQ	F 2. 0	84- 85
HELPTK	F 2. 0	86- 87

HELPTU	F 2. 0	88- 89
HELPAUTO	F 2. 0	90- 91
SCREENS	F 3. 0	92- 94
SEARCHES	F 3. 0	95- 97
DISPTOT	F 3. 0	98- 100
DISPEXP	F 3. 0	101- 103
DISPCONT	F 3. 0	104- 106
ERRORS	F 3. 0	107- 109
BROWSE	F 3. 0	110- 112

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. SESSION: This is the number assigned by the catalog to a particular user at logon.

3. TERMLOC: This variable contains the hexadecimal port address of the terminal associated with a particular session. Since each hard-wired terminal on the campuses has a fixed port address, this variable will allow us to select sessions based on campus location of the terminal. (Eventually two additional variables will be added to each record giving the WYLBUR USERID and GROUPID of the session - these are, again, fixed for each hardwired terminal.)

4. USERTYPE: USERTYPE currently is not filled in by SABS. It is intended to contain a code indicating the user's status (e.g., student, faculty, librarian, etc.) when that information becomes available through on-line questionnaires.

5. SRCHMODE: This variable contains a code to indicate the searching mode set for a particular user at the time the record was generated (e.g., LOOKUP, COMMAND, or not yet set).

6. STARTHR, STARTMIN & STARTSEC: These three variables record the clock time that the session started.

7. ENDHR, ENDMIN, & ENDSEC: These three variables record the clock time that the session ended.

8. LENGTH: This variable records the session length in second and fractions of a second.

9. ADACPU: Total ADABAS CPU time consumed by a particular session (not currently filled in).

10. ADATOTX: Total ADABAS EXCPs performed during the session (not currently filled in).

11. ADADATX: ADABAS data storage EXCPs performed during the session (not currently filled in).

12. ADAASSX: ADABAS associator EXCPs performed during the session (not currently filled in).

13. ADAWRKX: ADABAS work storage EXCPs performed during the session (not currently filled in).

14. COMMANDS: Total number of commands of all types issued during the session.

15. BREAKS: Number of times break was hit during the session (not currently logged except as total breaks during an interval in NCDQ).

16. PARMTOT: Number of times session parameters were set or changed by the user.

17. PARMTRM: The last or default value set for terminal type for the session.

18. PARMFRM: The last or default value set for display format for the session.

19. PARMLIB: The last or default value set for location (library) for the session.

20. PARMLAN: The last or default value set for language for the session.

21. HELPTOT: Total number of times HELP was evoked for the session.

22. HELPUNQ: Number of times a HELP command was issued alone (with no glossary terms).

23. HELPTK: Number of times a help command was issued with a term, and the term was found in the glossary.

24. HELPTU: Number of times a help command was issued with a term, and the term was not found in the glossary.

25. HELPAUTO: Number of times that QHELP automatically sent a help screen to a user (for three identical consecutive errors).

26. SCREENS: Total number of 'canned' screens sent to the user during the session.

27. SEARCHES: Number of valid (syntactically correct) searches made by the user during the session.

28. DISPTOT: Total number of displays sent to the user (a display = 1 display screen) during the session.

29. DISPEXP: Number of displays evoked by explicit display commands.

30. DISPCONT: Number of times a display was continued with <cr> or NS or PS.

31. ERRORS: Total number of errors detected during the session.

32. BROWSE: Total number of syntactically correct BROWSE commands issued during the session.

C.4 NCDT - SEARCH SUMMARY FILE

VARIABLE	FORMAT	COLUMNS
YEAR	A 2	1- 2
MONTH	A 2	3- 4
DAY	A 2	5- 6
SESSION	F 6. 0	7- 12
LOGNO	F 4. 0	13- 16
USERTYPE	A 2	17- 18
SRCHMODE	A 1	19- 19
SCHTYPE1	A 1	20- 20
INDICES	F 2. 0	21- 22
SCHTYPE2	A 1	23- 23
SCHTYPE3	A 1	24- 24
PRIORITY	A 1	25- 25
RETRTYPE	A 1	26- 26
RETRIEV	F 4. 0	27- 30
RECS_IN	F 4. 0	31- 34
LOCAT	A 1	35- 35
LOCRTR	F 4. 0	36- 39
TERMINAT	A 1	40- 40
ERRCODE	F 4. 0	41- 44
SRCHLEN	F10. 3	45- 54
ADACPU	F 9. 2	55- 63
ADATOTX	F 4. 0	64- 67
ADADATX	F 3. 0	68- 70
ADAASSX	F 3. 0	71- 73
ADAWRKX	F 3. 0	74- 76
INDEX1	A 2	77- 78
INDEX2	A 2	79- 80
INDEX3	A 2	81- 82
INDEX4	A 2	83- 84
INDEX5	A 2	85- 86
INDEX6	A 2	87- 88

OPER1	A	1	89- 89
OPER2	A	1	90- 90
OPER3	A	1	91- 91
OPER4	A	1	92- 92
OPER5	A	1	93- 93
OPER6	A	1	94- 94
KEY1	A	1	95- 95
MOD1	A	1	96- 96
KEY2	A	1	97- 97
MOD2	A	1	98- 98
KEY3	A	1	99- 99
MOD3	A	1	100- 100
KEY4	A	1	101- 101
MOD4	A	1	102- 102
KEY5	A	1	103- 103
MOD5	A	1	104- 104
KEY6	A	1	105- 105
MOD6	A	1	106- 106

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. SESSION: This is the number assigned by the catalog to a particular user at logon.

3. LOGNO: This is the sequence number within a particular session of the NFAC record providing the basis for each record of this file. This variable is included to aid in linking records from different files that relate to the same event.

4. USERTYPE: USERTYPE currently is not filled in by SABS. It is intended to contain a code indicating the user's status (e.g., student, faculty, librarian, etc.) when that information becomes available through on-line questionnaires.

5. SRCHMODE: This variable contains a code to indicate the searching mode set for a particular user at the time the record was generated. (E.g., LOOKUP, COMMAND, or not yet set.)

6. SCHTYPE1: Contains code for type of search command--initiating (i.e., FIND), modifying (i.e., AND, OR, AND NOT), location processing (i.e., AT), or selection from a Browse (i.e., SELECT 1-3).

7. INDICES: The number of indices specified in the search statement.

8. SCHTYPE2: Types of indices searched for this search (primary, secondary, or location.)

9. SCHTYPE3: Search execution type: normal or long search.

10. PRIORITY: Priority set for this search (HIGH, LOW, or NORMAL) (priorities are not currently set by the patron interface).

11. RETRTYPE: Type of records retrieved (headings or books).

12. RETRIEV: Number of items retrieved by the search (systemwide count).

13. RECS_IN: For modifying searches, the number of items in the previously retrieved set.

14. LOCATION: Location code in effect for the search (i.e., library set).

15. LOCRTR: Number of records of the retrieved set at the specified location.

16. TERMINAT: Contains a code for the way the search ended (N = normal, Q = canceled long search, B = break key hit).

17. ERRCODE: Not currently filled in: search statements with errors are logged as error text records and error summary records (see NCDW and NFAP).

18. SRCHLEN: Elapsed time in seconds and fractions of a second from the time the user hit return on the search statement, to the time the search was completed and the results logged (just prior to sending the retrieval report to the user).

19. ADACPU: ADABAS CPU time consumed by the search command processing (not currently filled in).

20. ADATOTX: Total ADABAS EXCPs used in execution of the search command (not currently filled in).

21. ADADATX: ADABAS data storage EXCPs used in execution of the search command (not currently filled in).

22. ADAASSX: ADABAS associator EXCPs used in execution of the search command (not currently filled in).

23. ADAWRKX: ADABAS work storage EXCPs used in execution of the search command (not currently filled in).

24. INDEX1 to INDEX6: In sequence, the indexes used in the search command.

25. OPER1 to :OPER6: In sequence, the Boolean operators used to connect indexes.

26. KEY1 to KEY6 and MOD1 to MOD6: Types (or numbers) or keywords specified for each index (e.g., single word, multiple word, multiple word with Boolean operators between keywords, etc.) also, modifications to keywords, (e.g. truncation, Boolean operators used between keywords, etc.).

C.5 NCDU - COMMAND TRANSITION FILE

VARIABLE	FORMAT	COLUMNS
YEAR	A 2	1- 2
MONTH	A 2	3- 4
DAY	A 2	5- 6
HOUR	F 2. 0	7- 8
MIN	F 2. 0	9- 10
SEC	F 2. 0	11- 12
SESSION	F 6. 0	13- 18
LOGNO	F 4. 0	19- 22
SRCHMODE	A 1	23- 23
PREVCOM	A 4	24- 27
NEWCOM	A 4	28- 31
TRNSTIME	F10. 3	32- 41

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. HOUR, MIN, and SEC: These three variables contain the time of day that the NFAC record providing a basis for this record was generated. The variables are included to allow aggregation of records by time of day in SPSS processing.

3. SESSION: This is the number assigned by the catalog to a particular user at logon.

4. LOGNO: This is the sequence number within a particular session of the NFAC record providing the basis

for each record of this file. This variable is included to aid in linking records from different files that relate to the same event.

5. SRCHMODE: This variable contains a code to indicate the searching mode set for a particular user at the time the record was generated. (E.g., LOOKUP, COMMAND, or not yet set.)

6. PREVCOM: Contains code value representing the previous command type issued in a user's session.

7. NEWCOM: Contains a code value representing the 'current' command type issued in a user's session.

8. TRNSTIME: The elapsed time in seconds and fractions of a second between the previous command and the current command.

C.6 NCDV - SEARCH TEXT FILE

VARIABLE	FORMAT	COLUMNS
YEAR	A 2	1- 2
MONTH	A 2	3- 4
DAY	A 2	5- 6
HOUR	F 2. 0	7- 8
MIN	F 2. 0	9- 10
SEC	F 2. 0	11- 12
SESSION	F 6. 0	13- 18
LOGNO	F 4. 0	19- 22
SRCHMODE	A 1	23- 23
RETRIEVE	F 4. 0	24- 27
STATCODE	F 3. 0	28- 30
PCTLEN	F 3. 0	31- 33
COMMLEN	B 4	34- 37
COMMAND	A 150	38- 188

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. HOUR, MINUTE, and SECOND: These three variables contain the time of day that the NFAC record providing a basis for this record was generated. The variables are included to allow aggregation of records by time of day in SPSS processing.

3. SESSION: This is the number assigned by the catalog to a particular user at logon.

4. LOGNO: This is the sequence number within a particular session of the NFAC record providing the basis for each record of this file. This variable is included to aid in linking records from different files that relate to the same event.

5. SRCHMODE: This variable contains a code to indicate the searching mode set for a particular user at the time the record was generated. (E.g., LOOKUP, COMMAND, or not yet set.)

6. RETRIEVE: Number of items retrieved by the logged command.

7. STATCODE: '000' for normal searches, '999' for long searches.

8. PCTLEN: Length of the command string (human and SPSS readable).

9. COMMLEN: Length of the command string (in full-word binary format -- machine-readable only).

10. COMMAND: Text of the search command.

C.7 NCDW - ERROR TEXT FILE

VARIABLE	FORMAT	COLUMNS
YEAR	A 2	1- 2
MONTH	A 2	3- 4
DAY	A 2	5- 6
HOUR	F 2. 0	7- 8
MIN	F 2. 0	9- 10
SEC	F 2. 0	11- 12
SESSION	F 6. 0	13- 18
LOGNO	F 4. 0	19- 22
SRCHMODE	A 1	23- 23
RETRIEVE	F 4. 0	24- 27
STATCODE	F 3. 0	28- 30
PCTLEN	F 3. 0	31- 33
COMMLEN	B 4	34- 37
COMMAND	A 150	38- 188

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. HOUR, MINUTE, and SECOND: These three variables contain the time of day that the NFAC record providing a basis for this record was generated. The variables are included to allow aggregation of records by time of day in SPSS processing.

3. SESSION: This is the number assigned by the catalog to a particular user at logon.

4. LOGNO: This is the sequence number within a particular session of the NFAC record providing the basis for each record of this file. This variable is included to aid in linking records from different files that relate to the same event.

5. SRCHMODE: This variable contains a code to indicate the searching mode set for a particular user at the time the record was generated. (E.g., LOOKUP, COMMAND, or not yet set.)

6. RETRIEVE: Number of items retrieved by the logged command (always 0 for errors).

7. STATCODE: The user session status as reported by QHELP when the error occurred. (The status codes indicate the most recent events in the user session.)

8. PCTLEN: Length of the command string (human and SPSS readable).

9. COMMLEN: Length of the command string (in full-word binary format -- machine-readable only).

10. COMMAND: Text of the command causing an error.

C.8 NCDX - SCREEN TRANSITION FILE

VARIABLE	FORMAT	COLUMNS
YEAR	A 2	1- 2
MONTH	A 2	3- 4
DAY	A 2	5- 6
HOUR	F 2. 0	7- 8
MIN	F 2. 0	9- 10
SEC	F 2. 0	11- 12
SESSION	F 6. 0	13- 18
SRCHMODE	A 1	19- 19
LOGNO	F 4. 0	20- 23
PRVSCRN	F 7. 0	24- 30
NEWSCRN	F 7. 0	31- 37
ORIGIN	F 1. 0	38- 38
TRNSTIME	F10. 3	39- 48

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. HOUR, MINUTE, and SECOND: These three variables contain the time of day that the NFAC record providing a basis for this record was generated. The variables are included to allow aggregation of records by time of day in SPSS processing.

3. SESSION: This is the number assigned by the catalog to a particular user at logon.

4. SRCHMODE: This variable contains a code to indicate the searching mode set for a particular user at the time the record was generated. (E.g., LOOKUP, COMMAND, or not yet set.)

5. LOGNO: This is the sequence number within a particular session of the NFAC record providing the basis for each record of this file. This variable is included to aid in linking records from different files that relate to the same event.

6. PRVSCRN: Previous screen sent to user terminal.

7. NEWSCRN: New (or current) screen sent to user terminal.

8. ORIGIN: Where the new screen came from before being sent to the user. '0' = read from screen file, '1' = read from in-core buffer.

9. TRNSTIME: The elapsed time in seconds and fractions of a second between the previous screen and current screen.

C.9 NCDY - SUBSYSTEM TRANSITION FILE

NCDY has been 'turned off' and records are not being produced. When 'turned on' NCDY generates records for each transition between subsystems (as represented by NFAC records). It creates a big file with little information.

C.10 NCDZ - SELECTIVE RECORD FILE

NCDZ is output in the same format as the NFAC input records. (See Appendix B for a description of NFAC records.) Records are written to this file based on criteria of 'peculiarity' (e.g., NFAC records generated by the HELP subsystem that have the verb 'HELP' followed by over 30 characters of text). The NCDZ file may also be used to write out special selections of NFAC records based on criteria input as control cards in a SCAN mode (yet to be implemented).

C.11 NFAO - HELP SUMMARY FILE

VARIABLE	FORMAT	COLUMNS
YEAR	A 2	1- 2
MONTH	A 2	3- 4
DAY	A 2	5- 6
SESSION	F 6. 0	7- 12
LOGNO	F 4. 0	13- 16
USERTYPE	A 2	17- 18
HELPCODE	F 1. 0	19- 19
STATCODE	F 3. 0	20- 22
WORD1	A 4	23- 26
WORD2	A 4	27- 30
WORD3	A 4	31- 34
WORD4	A 4	35- 38
WORD5	A 4	39- 42
WORD6	A 4	43- 46
WORD7	A 4	47- 50
REST	A 2	51- 52

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. SESSION: This is the number assigned by the catalog to a particular user at logon.

3. LOGNO: This is the sequence number within a particular session of the NFAC record providing the basis for each record of this file. This variable is included to aid in linking records from different files that relate to the same event.

4. USERTYPE: USERTYPE currently is not filled in by SABS. It is intended to contain a code indicating the

user's status (e.g., student, faculty, librarian, etc.) when that information becomes available through on-line questionnaires.

5. HELPCODE: Type of help requested: '1' Unqualified, '2' with known term, or '3' with unknown term.

6. STATCODE: User's status at the time help was requested. (The status codes indicate the most recent events in the user session.)

7. WORD1 to WORD7 & REST: These variables actually represent a 30-character string which contains the word (or words) following the verb in a HELP command (if any). If more than 30 characters occur, the NFAC record is written to NCDZ. The string is chopped up into segments of 4 characters to make it digestible to SPSS. For actual frequency counts on the words contained in this text section, the same records may be processed by the concordance programs.

C.12 NFAP - ERROR SUMMARY FILE

VARIABLE	FORMAT	COLUMNS
YEAR	A 2	1- 2
MONTH	A 2	3- 4
DAY	A 2	5- 6
SESSION	F 6. 0	7- 12
LOGNO	F 4. 0	13- 16
USERTYPE	A 2	17- 18
HELPCODE	F 1. 0	19- 19
STATCODE	F 3. 0	20- 22

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. SESSION: This is the number assigned by the catalog to a particular user at logon.

3. LOGNO: This is the sequence number within a particular session of the NFAC record providing the basis for each record of this file. This variable is included to aid in linking records from different files that relate to the same event.

4. USERTYPE: USERTYPE currently is not filled in by SABS. It is intended to contain a code indicating the user's status (e.g., student, faculty, librarian, etc.) when that information becomes available through on-line questionnaires.

5. HELPCODE: Type response to error: '4' Short error message sent to user, or '5' Same error repeated three times--automatic help screen sent out.

6. STATCODE: User's status at the time the error occurred. (The status codes indicate the most recent events in the user session.)

C.13 NFAQ - BROWSE SUMMARY FILE

VARIABLE	FORMAT	COLUMNS
YEAR	A 2	1- 2
MONTH	A 2	3- 4
DAY	A 2	5- 6
SESSION	F 6. 0	7- 12
LOGNO	F 4. 0	13- 16
USERTYPE	A 2	17- 18
RESPTIME	F10. 3	19- 28
INDEX	A 2	51- 52
KEYTYPE	A 1	53- 53
KEYMOD	A 1	54- 54

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. SESSION: This is the number assigned by the catalog to a particular user at logon.

3. LOGNO: This is the sequence number within a particular session of the NFAC record providing the basis for each record of this file. This variable is included to aid in linking records from different files that relate to the same event.

4. USERTYPE: USERTYPE currently is not filled in by SABS. It is intended to contain a code indicating the user's status (e.g., student, faculty, librarian, etc.) when that information becomes available through on-line questionnaires.

5. RESPTIME: Elapsed time in seconds and fractions of a second from the time return was pressed on a Browse command to the time the record was logged (just prior to displaying the first screen of the browse.)

6. INDEX: The primary index which was 'browsed.'

7. KEYTYPE: Contains codes to indicate single, multiple, or multiple Boolean keywords used in the browse command.

8. KEYMOD: Contains codes for any modifications to the keywords (truncation, Boolean operators, phrase searching, etc.).

C.14 NFAR - BROWSE TEXT FILE

VARIABLE	FORMAT	COLUMNS
YEAR	A 2	1- 2
MONTH	A 2	3- 4
DAY	A 2	5- 6
HOUR	F 2. 0	7- 8
MIN	F 2. 0	9- 10
SEC	F 2. 0	11- 12
SESSION	F 6. 0	13- 18
LOGNO	F 4. 0	19- 22
SRCHMODE	A 1	23- 23
RETRIEVE	F 4. 0	24- 27
STATCODE	F 3. 0	28- 30
PCTLEN	F 3. 0	31- 33
COMMLEN	B 4	34- 37
COMMAND .	A 150'	38- 188

1. YEAR, MONTH, and DAY: These three variables contain the date that the NFAC record providing the basis for this record was generated. The variables are included to allow aggregation of records by date in SPSS processing when records from several days are combined in one file.

2. HOUR, MINUTE, and SECOND: These three variables contain the time of day that the NFAC record providing a basis for this record was generated. The variables are included to allow aggregation of records by time of day in SPSS processing.

3. SESSION: This is the number assigned by the catalog to a particular user at logon.

4. LOGNO: This is the sequence number within a particular session of the NFAC record providing the basis for each record of this file. This variable is included to aid in linking records from different files that relate to the same event.

5. SRCHMODE: This variable contains a code to indicate the searching mode set for a particular user at the time the record was generated. (E.g., LOOKUP, COMMAND, or not yet set.)

6. RETRIEVE: Number of items retrieved by the logged command.

7. STATCODE: Always blank or zero for browses.

8. PCTLEN: Length of the command string (human and SPSS readable).

9. COMMLEN: Length of the command string (in full-word binary format -- machine-readable only).

10. COMMAND: Text of the browse command.

Appendix D

SPSS PROGRAMS

This appendix describes the SPSS programs currently available to process the files created by SABS of statistical data from the prototype on-line catalog.

D.1 ACTSUM : ACTIVITY SUMMARY PROCESSING

SPSS program ACTSUM uses NCDQ - the activity summary file from SABS TIME SCAN runs - to provide measures of central tendency for all pertinent variables in that file. What emerges is a picture of an 'average' 5 minute slice of catalog activity for a given time frame (i.e., if the file being processed is a daily file, it shows an average 'slice for that day', if the file is a weekly file - for that week, etc.). For each variable the mean, variance, range, sum, standard error, standard deviation, kurtosis, skewness, minimum value, and maximum value are reported.

D.2 BRWSREC : BROWSE SUMMARY PROCESSING

SPSS program BRWSREC uses the NFAQ file - browse summary records - from SABS NORM runs. Currently BRWSREC computes average response time for all browse records in the file. It also provides frequency breakdowns of the indexes browsed, the types of keywords employed, and the modifications to keywords.

D.3 COMMTRNS : COMMAND TRANSITION PROCESSING

SPSS program COMMTRNS uses the NCDU file (command transitions) produced by SABS NORM runs. The program discriminates between commands issued in LOOKUP mode sessions and COMMAND mode sessions. For each mode, frequencies are provided for the last and current command types issued in a session. Average transition time between commands is calculated and reported. Also for each mode a transition matrix showing the frequency of transition from one command type to another is generated.

D.4 DISPSUM : DISPLAY SUMMARY PROCESSING

SPSS program DISPSUM uses the NCDR file (display summary records) produced by SABS NORM runs. The program discriminates between records from LOOKUP and COMMAND mode sessions. For each search mode DISPSUM computes the measures of central tendency for variables DISPTIME, USERTIME, and NFORM (see the file description of NCDR). It also produces frequency tables for FORMAT2, FORMAT3, TERMIN, RANGE, CURREC, RECCNT, SCREEN, and PART (for each mode).

D.5 ERRREC : ERROR SUMMARY PROCESSING

SPSS program ERRREC uses the NFAP file (error summary records) produced by SABS NORM mode runs. It generates four frequency tables;

1. Frequency of errors by response (accually a count of the records that will appear in 3 and 4).
2. Frequency of error types overall.
3. Frequency of errors generating a short error message.
4. Frequency of errors causing an automatic help screen response (when three identical sequential errors occur).

D.6 HELPREC : HELP SUMMARY PROCESSING

SPSS program HELPREC uses file NFAO (help summary records) produced by SABS NORM mode runs. It generates frequency tables of the type of help requested (unqualified, with known glossary word, or with unknown word) and the session status at the time help was requested for each type. It also produces a frequency table of the first four characters of words following the 'HELP' verb. (Eventually this will be done better by a concordance program - but it does give a good indication of which glossary words are being asked for.)

D.7 NEWADA : ADABAS LOG FILE PROCESSING

SPSS program NEWADA uses the log file generated by ADABAS. It generates frequency tables for all variables in the ADABAS log file. (See section 2.2.2 for variable descriptions.) CPU time is recoded to a set of ranges and EXCPs are totaled for the frequency tables. Cross-tabulations are performed to show the CPU times and EXCPs associated with each file and with each command type.

D.8 SCRNTRNS : SCREEN TRANSITION PROCESSING

SPSS program SCRNTRNS uses NCDX (screen transition file) produced by SABS NORM runs. It generates separate frequency tables for screens transmitted to users in LOOKUP and COMMAND mode sessions. It also provides measures of central tendency for the TRNSTIME variable (elapsed time between screens).

D.9 SESSUM : SESSION SUMMARY PROCESSING

SPSS program SESSUM use the NCDS file (session summary records) produced by SABS NORM runs. It produces frequency tables for variables SRCHMODE, ENDTYPE, TERMLOC, PARMTRM, PARMFRM, PARMLIB, and PARMLAN. It computes measures of central tendency for variables LENGTH, COMMANDS, BREAKS, PARMTOT, HELPTOT, HELPUNQ, HELPTK, HELPTU, HELPAUTO, SCREENS, SEARCHES, DISPTOT, DISPEXP, DISPCONT, ERRORS, and BROWSE. (See the description of NCDS in Appendix C for further description of these variables).

D.10 SRCHSUM : SEARCH SUMMARY PROCESSING

SPSS program SRCHSUM uses the NCDT file (search summary records) produced by SABS NORM runs. For both LOOKUP and COMMAND mode records, frequency tables are generated for each of the following variables: SRCHMODE, SCHTYPE1, INDICES, SCHTYPE2, SCHTYPE3, TERMINAT, LOCAT, INDEX1 TO INDEX6, OPER1 TO OPER6, KEY1 TO KEY6, AND MOD1 TO MOD6. Some cross-tabulations between indexes and keyword types are done. Measures of central tendency are computed for SRCHLEN (response time), RETRIEV (number of items retrieved), LOCRTR, and RECS_IN.

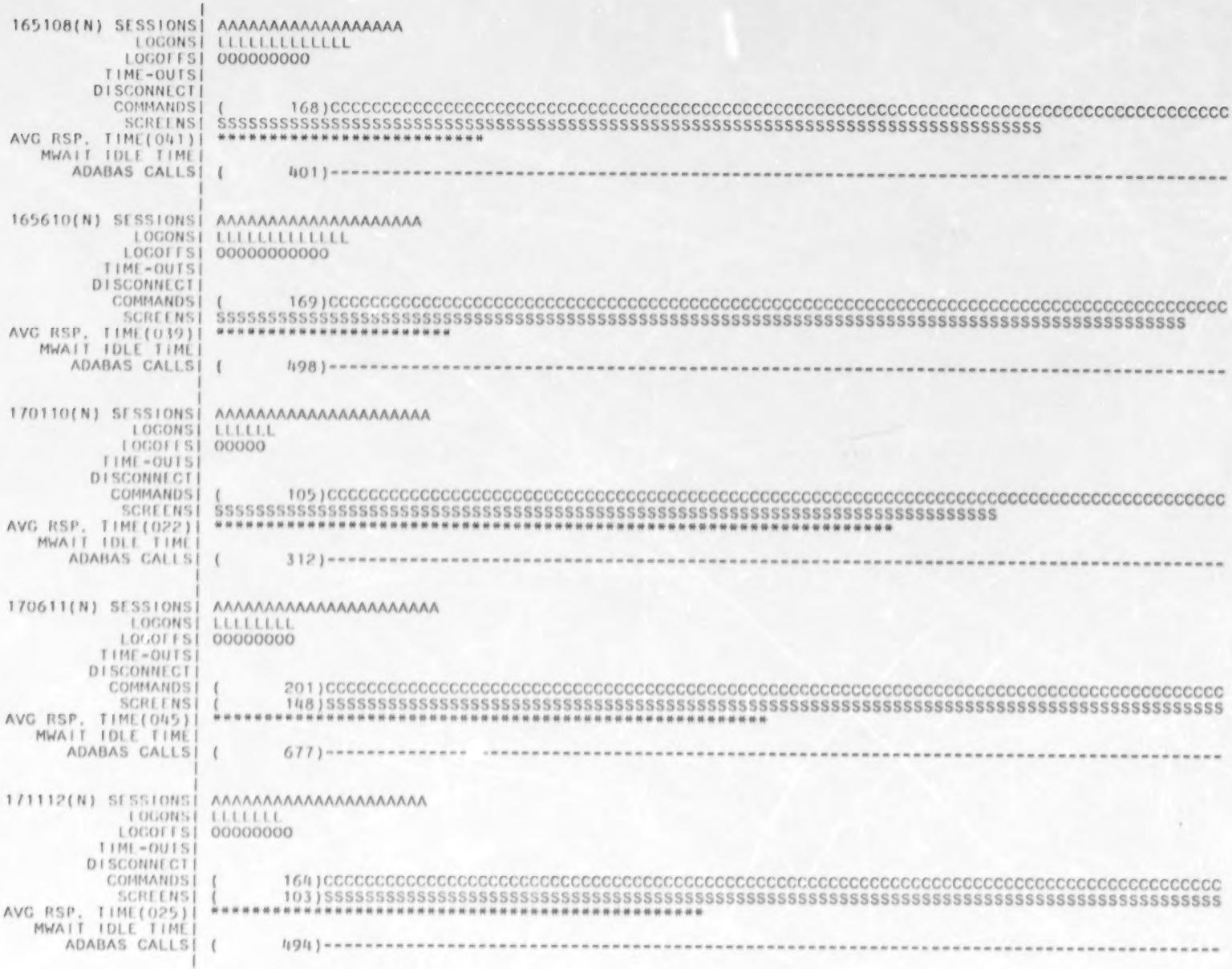
Appendix E

REPORT EXAMPLES

The following pages contain examples of the reports produced by SABS and SPSS processing.

```
*****************************************************************************************************************************
*                                                                                                                           *
*                  HISTOGRAM OF TIME SCAN        INTERVAL = 300      NFAC DATE: 810727   SYSTEM START TIME: 162554      *
*                                                                                                                           *
*****************************************************************************************************************************

                                1    1    2    2    3    3    4    4    5    5    6    6    7    7    8    8    9    9
                   012345678901234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
                  ++----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
                   |
                   |
                   |
163057(N) SESSIONS| AAAAAAAAAA
             LOGONS| LLLLLLLLLLL
            LOGOFFS| OO
          TIME-OUTS|
         DISCONNECT|
           COMMANDS| CCCC
```


165108(N) SESSIONS
LOGONS
LOGOFFS
TIME-OUTS
DISCONNECT
COMMANDS
SCREENS
AVG RSP. TIME(041)
MWAIT IDLE TIME
ADABAS CALLS
165610(N) SESSIONS
LOGONS
LOGOFFS
TIME-OUTS
DISCONNECT
COMMANDS
SCREENS
AVG RSP. TIME(039)
MWAIT IDLE TIME
ADABAS CALLS
170110(N) SESSIONS
LOGONS
LOGOFFS
TIME-OUTS
DISCONNECT
COMMANDS
SCREENS
AVG RSP. TIME(022)
MWAIT IDLE TIME
ADABAS CALLS
170611(N) SESSIONS
LOGONS
LOGOFFS
TIME-OUTS
DISCONNECT
COMMANDS
SCREENS
AVG RSP. TIME(045)
MWAIT IDLE TIME
ADABAS CALLS
171112(N) SESSIONS
LOGONS
LOGOFFS
TIME-OUTS
DISCONNECT
COMMANDS
SCREENS
AVG RSP. TIME(025)
MWAIT IDLE TIME
ADABAS CALLS

FILE NCDQTST (CREATION DATE = 07/29/81)

VARIABLE IDLE MWAIT IDLE TIME DURING INTERVAL IN SECON

MEAN	185.610	STD ERROR	6.963	STD DEV	113.561
VARIANCE	12895.994	KURTOSIS	-1.506	SKEWNESS	-0.081
RANGE	508.860	MINIMUM	0.501	MAXIMUM	509.361
SUM	49372.243				

VALID OBSERVATIONS - 266 MISSING OBSERVATIONS - 46

- - - -

FILE NCDQTST (CREATION DATE = 07/29/81)

VARIABLE SCRZENS NUMBER OF SCREENS TRANSMITTED DURING INT

MEAN	86.310	STD ERROR	2.999	STD DEV	42.734
VARIANCE	1826.235	KURTOSIS	-0.324	SKEWNESS	-0.177
RANGE	187.000	MINIMUM	1.000	MAXIMUM	188.000
SUM	17521.000				

VALID OBSERVATIONS - 203 MISSING OBSERVATIONS - 109

- - - -

FILE NCDQTST (CREATION DATE = 07/29/81)

VARIABLE TIMEOUTT NUMBER OF TIMEOUTS DURING INTERVAL

STATISTICS CAN NOT BE COMPUTED FOR THIS VARIABLE.
VARIABLE IS EITHER MISSING FOR EVERY CASE, ALPHANUMERIC, OR HAS NUMERIC VALUES EXCEEDING 10,000,000,000,000.

- - - -

FILE NCDSIST (CREATION DATE = 07/29/81)

SRCHMODE SEARCHING MODE SET FOR SESSION

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
COMMAND MODE	C	362	33.3	33.3	33.3
LOOKUP MODE	L	573	52.7	52.7	86.0
MODE NOT SET	X	152	14.0	14.0	100.0
	TOTAL	1087	100.0	100.0	

VALID CASES 1087 MISSING CASES 0

FILE NCDSTST (CREATION DATE = 07/29/81)

VARIABLE LENGTH SESSION LENGTH IN SECONDS

MEAN	600.461	STD ERROR	33.027	STD DEV	1079.336
VARIANCE	**********	KURTOSIS	39.677	SKEWNESS	5.322
RANGE	12285.961	MINIMUM	3.734	MAXIMUM	12289.695
SUM	641292.833				

VALID OBSERVATIONS - 1068 MISSING OBSERVATIONS - 19

- - - -

FILE NCDSTST (CREATION DATE = 07/29/81)

VARIABLE HELPUNQ NUMBER OF UNQUALIFIED HELP REQUESTS

MEAN	1.212	STD ERROR	0.038	STD DEV	0.501
VARIANCE	0.251	KURTOSIS	7.797	SKEWNESS	2.645
RANGE	3.000	MINIMUM	1.000	MAXIMUM	4.000
SUM	206.000				

VALID OBSERVATIONS - 170 MISSING OBSERVATIONS - 917

- - - -

FILE NCDSIST (CREATION DATE = 07/29/81)

VARIABLE SCREENS NUMBER OF SCREENS SENT TO TERMINAL

MEAN	11.990	STD ERROR	0.570	STD DEV	18.788
VARIANCE	352.999	KURTOSIS	38.865	SKEWNESS	4.973
RANGE	245.000	MINIMUM	1.000	MAXIMUM	246.000
SUM	13033.000				

VALID OBSERVATIONS - 1087 MISSING OBSERVATIONS - 0

- - - -

FILE NCDSTST (CREATION DATE = 07/29/81)

VARIABLE DISPCONT DISPLAY CONTINUATIONS -PS NS RETURN-

MEAN	11.952	STD ERROR	1.120	STD DEV	16.912
VARIANCE	286.002	KURTOSIS	12.571	SKEWNESS	3.178
RANGE	108.000	MINIMUM	1.000	MAXIMUM	109.000
SUM	2725.001				

VALID OBSERVATIONS - 228 MISSING OBSERVATIONS - 859

- - - -

FILE NCDTTST (CREATION DATE = 07/29/81)

SRCHMODE SEARCH MODE SET BY USER OR DEFAULT

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
COMMAND	C	2148	52.2	53.0	53.0
LOOKUP	L	1904	46.3	47.0	100.0
		63	1.5	MISSING	100.0
	TOTAL	4115	100.0	100.0	

VALID CASES 4052 MISSING CASES 63

FILE NCDTTST (CREATION DATE = 07/29/81)

INDEX1 1ST SEARCH INDEX

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
CORPORATE AUTHOR	CA	42	2.6	2.6	2.6
DATE OF PUBL.	DA	29	1.8	1.8	4.3
PERSONAL AUTHOR	PA	446	27.1	27.1	31.4
SERIES	SE	50	3.0	3.0	34.4
SUBJECT	SU	568	34.5	34.5	69.0
TITLE	TI	508	30.9	30.9	99.8
UNIFORM TITLE	UT	3	0.2	0.2	100.0
	TOTAL	1646	100.0	100.0	

VALID CASES 1646 MISSING CASES 0

FILE NONAME (CREATION DATE = 07/29/81)

DESCRIPTION OF SUBPOPULATIONS

CRITERION VARIABLE SRCHLEN SEARCH RESPONSE TIME IN SECONDS
BROKEN DOWN BY INDEX1 1ST SEARCH INDEX
BY KEY1 1ST KEYWORD TYPE

VARIABLE	CODE	VALUE LABEL	SUM	MEAN	STD DEV	VARIANCE	N
FOR ENTIRE POPULATION			6003.3641	3.7615	4.4331	19.6521	(1596)
INDEX1	UT	UNIFORM TITLE	4.7610	1.5870	0.9108	0.8296	(3)
KEY1	S	SINGLE WORD	2.1470	2.1470	0.0	0.0	(1)
KEY1	M	DEFAULT MULTI WORD	2.6140	1.3070	1.0904	1.1889	(2)
INDEX1	TI	TITLE	2628.7536	5.1747	4.3762	19.1509	(508)
KEY1	S	SINGLE WORD	444.7099	4.4031	4.6645	21.7576	(101)
KEY1	M	DEFAULT MULTI WORD	1992.0787	5.2981	4.2805	18.3224	(376)
KEY1	B	BOOLEAN MULTI WORD	191.9650	6.1924	4.3392	18.8289	(31)
INDEX1	SU	SUBJECT	2077.8527	3.6582	4.4310	19.6335	(568)
KEY1	S	SINGLE WORD	767.6739	4.0834	5.6492	31.9135	(188)
KEY1	M	DEFAULT MULTI WORD	1208.5748	3.3665	3.6540	13.3515	(359)
KEY1	B	BOOLEAN MULTI WORD	101.6040	4.8383	3.8130	14.5392	(21)
INDEX1	PA	PERSONAL AUTHOR	893.6478	2.0037	1.7329	3.0030	(446)
KEY1	S	SINGLE WORD	228.4140	2.5379	2.1347	4.5568	(90)
KEY1	M	DEFAULT MULTI WORD	652.1819	1.8687	1.5981	2.5540	(349)
KEY1	B	BOOLEAN MULTI WORD	13.0520	1.8646	1.2781	1.6336	(7)
INDEX1	DA	DATE OF PUBL.	290.0740	10.0026	13.8832	192.7440	(29)
KEY1	S	SINGLE WORD	290.0740	10.0026	13.8832	192.7440	(29)
INDEX1	CA	CORPORATE AUTHOR	108.2750	2.5780	2.0195	4.0786	(42)
KEY1	S	SINGLE WORD	13.8130	2.3022	2.2443	5.0367	(6)
KEY1	M	DEFAULT MULTI WORD	94.4620	2.6239	2.0107	4.0430	(36)

TOTAL CASES = 1596

FILE NCDTTST (CREATION DATE = 07/29/81)

* CROSSTABULATION OF *
INDEX1 1ST SEARCH INDEX BY INDEX2 2ND SEARCH INDEX
* * * * * * * * * * * * * * * * * * * \

FILE NCDTTST (CREATION DATE = 07/29/81)

* CROSSTABULATION OF *
INDEX1 1ST SEARCH INDEX BY INDEX2 2ND SEARCH INDEX
* * * * * * * * * * * * * * * * * * * \

FILE NCDRTST (CREATION DATE = 07/29/81)

FORMAT3 FORMATTING USED IN THIS DISPLAY

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
BRIEF DISPLAY	0	2553	62.8	62.8	62.8
PARAGRAPH DISPLAY	1	685	16.8	16.8	79.6
MARC DISPLAY	2	115	2.8	2.8	82.5
REVIEW DISPLAY	3	327	8.0	8.0	90.5
HEADINGS DISPLAY	4	307	7.6	7.6	98.1
FIELD -TAG- DISPLAY	5	64	1.6	1.6	99.6
	6	15	0.4	0.4	100.0
	TOTAL	4066	100.0	100.0	

VALID CASES 4066 MISSING CASES 0

FILE NCDRTST (CREATION DATE = 07/29/81)

VARIABLE DISPTIME RESPONSE TIME FROM COMMAND TO DISPLAY

MEAN	0.718	STD ERROR	0.016	STD DEV	1.019
VARIANCE	1.038	KURTOSIS	105.603	SKEWNESS	7.346
RANGE	24.133	MINIMUM	0.002	MAXIMUM	24.135
SUM	2918.388				

VALID OBSERVATIONS - 4066 MISSING OBSERVATIONS - 0

- - - -

FILE NCDUTST (CREATION DATE = 07/29/81)

NEWCOM CURRENT COMMAND ISSUED IN SESSION

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
UNKNOWN VERB OR ERRO	????	3578	16.8	16.8	16.8
	BRER	40	0.2	0.2	17.0
BROWSE COMMAND	BRWS	209	1.0	1.0	18.0
DISPLAY COMM. ERROR	DERR	154	0.7	0.7	18.7
DISP. NEXT SCREEN	DSNS	185	0.9	0.9	19.6
DISP. PREV. SCREEN	DSPS	113	0.5	0.5	20.1
BRIEF DISPLAY REQ.	DSP0	570	2.7	2.7	22.8
LONG DISPLAY REQ.	DSP1	233	1.1	1.1	23.9
MARC DISPLAY REQ.	DSP2	28	0.1	0.1	24.0
REVIEW DISPLAY REQ.	DSP3	192	0.9	0.9	24.9
FIELDS DISPLAY REQ.	DSP5	33	0.2	0.2	25.1
	DSP6	10	0.0	0.0	25.2
HELP COMMAND	HELP	938	4.4	4.4	29.6
LOGOFF COMMAND	LGOF	405	1.9	1.9	31.5
	LKER	820	3.9	3.9	35.3
NUMBER	NUMB	7236	34.0	34.0	69.4
RESET LIBRARY	RESL	2	0.0	0.0	69.4
RESET TERMINAL	REST	2	0.0	0.0	69.4
CARRAIGE RETURN	RTRN	3729	17.5	17.5	86.9
SEARCH BACKUP	SCHB	8	0.0	0.0	87.0
INITIATING SEARCH	SCHI	1709	8.0	8.0	95.0
LOCATION PROCESSING	SCHL	251	1.2	1.2	96.2

FILE NCDUTST (CREATION DATE = 07/29/81)

MODIFYING SEARCH	SCHM	71	0.3	0.3	96.5
QUIT LONG SEARCH	SCHQ	10	0.0	0.0	96.6
SELECT FROM HEADINGS	SCHS	69	0.3	0.3	96.9
YES TO LONG SEARCH	SCHY	89	0.4	0.4	97.3
SEARCH COMM ERROR	SERR	229	1.1	1.1	98.4
SET FORMAT	SETF	8	0.0	0.0	98.4
SET LIBRARY	SETL	10	0.0	0.0	98.5
SET TERMINAL	SETT	2	0.0	0.0	98.5
SIGNON COMMAND	SIGN	313	1.5	1.5	100.0
	XXXX	6	0.0	0.0	100.0
	TOTAL	21252	100.0	100.0	

VALID CASES 21252 MISSING CASES 0

FILE NFAQTST (CREATION DATE = 07/29/81)

INDEX INDEX BROWSED

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
CORP. AUTHOR	CA	7	3.3	3.5	3.5
PERSONAL AUTHOR	PA	28	13.4	13.9	17.4
SERIES	SE	2	1.0	1.0	18.4
SUBJECT	SU	142	67.9	70.6	89.1
TITLE	TI	22	10.5	10.9	100.0
		8	3.8	MISSING	100.0
	TOTAL	209	100.0	100.0	

VALID CASES 201 MISSING CASES 8

FILE NFAQIST (CREATION DATE = 07/29/81)

KEYTYPE KEYWORD TYPE USED

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
BOOLEAN MULTI WORD	B	4	1.9	1.9	1.9
MULTI WORD	M	127	60.8	61.7	63.6
SINGLE WORD	S	75	35.9	36.4	100.0
		3	1.4	MISSING	100.0
	TOTAL	209	100.0	100.0	

VALID CASES 206 MISSING CASES 3

FILE NFAQTST (CREATION DATE = 07/29/81)

KEYMOD KEYWORD MODIFICATION OR BOOLEAN OPERATOR

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
BOOLEAN AND	A	4	1.9	10.8	10.8
BOOLEAN OR	O	2	1.0	5.4	16.2
TRUNCATION	T	26	12.4	70.3	86.5
AND + TRUNC.	1	5	2.4	13.5	100.0
		172	82.3	MISSING	100.0
	TOTAL	209	100.0	100.0	

VALID CASES 37 MISSING CASES 172